(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 April 2006 (20.04.2006)

PCT

(10) International Publication Number
WO 2006/042002 A2

(51) International Patent Classification:
A61K 38/17 (2006.01)

(21) International Application Number:
PCT/US2005/035961

(22) International Filing Date: 5 October 2005 (05.10.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/616,596	5 October 2004 (05.10.2004)	US
60/688,355	6 June 2005 (06.06.2005)	US

(71) Applicant *(for all designated States except US)*: **OREGON HEALTH AND SCIENCE UNIVERSITY** [US/US]; Technology and Research Collaborations, 2525 SW First Avenue, Suite 120, Mail Stop AD120, Portland, OR 97201 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CLINTON, Gail, M.** [US/US]; Oregon Health & Science University, 2525 SW First Avenue, Suite 120, Mail Stop AD120, Portland, OR 97301 (US). **ROBERTS, Charles, T., Jr.** [US/US]; Oregon Health & Science University, 2525 SW First Avenue, Suite 120, Mail Stop AD120, Portland, OR 97301 (US).

(74) Agents: **DAVISON, Barry, L.** et al.; Davis Wright Tremaine LLP, 2600 Century Square, 1501 Fourth Avenue, Seattle, WA 98101-1688 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

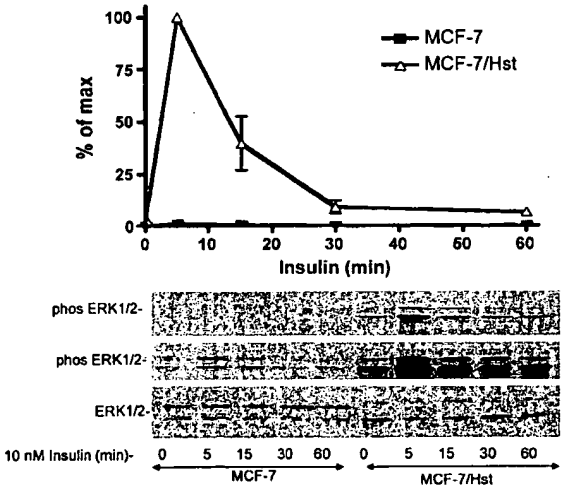
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR TREATING DISEASE



(57) Abstract: The present invention discloses for the first time that the insulin receptor (IR) is a target of Herstatin, which modulates IR and IR-mediated intracellular signaling. In preferred aspects, Herstatin binds at nM concentrations to cell-surface IR, up-regulates basal IR expression by several-fold, induces the accumulation of pro-IR, and stimulates insulin activation of the ERK pathway. Moreover, these changes in insulin signaling are accompanied by alterations in IGF-IR expression, IRS-2 levels, and the serine phosphorylation state of both IRS-1 and IRS-2. Preferred aspects provide novel therapeutic methods and pharmaceutical compositions for treatment of conditions associated with altered IR expression or IR-mediated signaling, including but not limited to insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof, and cancer.

WO 2006/042002 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

COMPOSITIONS AND METHODS FOR TREATING DISEASE

FIELD OF THE INVENTION

Aspects of the invention relate generally to therapeutic molecules, compositions and methods for treatment of diseases through modulation of the insulin receptor (IR) and IR-mediated intracellular signaling by administration of Herstatin or variants thereof, and in more particular aspects relate to compositions and methods for cell targeting, and for the treatment of conditions or diseases associated with altered IR expression or altered IR-mediated signaling, including but not limited to insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof, and cancer.

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of priority to: United States Provisional Patent Application Serial No. 60/616,596, filed 05 October 2004 and entitled "COMPOSITIONS AND METHODS FOR TREATING DISEASE"; and to United States Provisional Patent Application Serial No. 60/688,355, filed 06 June 2005, of same title, both of which are incorporated by reference herein in their entireties.

BACKGROUND

The Insulin Receptor. The insulin receptor is the canonical member of the insulin receptor family of receptor tyrosine kinases, which also includes the IGF-IR and the insulin receptor-related receptor (IRR). These molecules share a heterotetrameric structure comprised of two extracellular ligand-binding α subunits, which are coupled to each other and to two transmembrane β subunits by disulfide linkages. The intracellular portion of the β subunit contains the intrinsic tyrosine kinase catalytic domain, which is activated by binding of extracellular ligand and a presumed conformational change in the β subunit. The activated receptor undergoes autophosphorylation of tyrosine residues in the kinase domain as well as residues in the flanking juxtamembrane and carboxyl-terminal domains. The phosphorylation of these residues, particularly in the juxtamembrane region, allows the recruitment of scaffolding adapter proteins such as IRS-1 and IRS-2 and Shc, which are then phosphorylated on tyrosine residues by the activated receptor to recruit a second level of signaling molecules to initiate the signaling cascades that are responsible for insulin action. These include the ERK arm of the

MAPK pathway, the P13K-Akt/PKB pathway, and the APS-Cbl-CrkII-TC10 pathway. In cells expressing both insulin and IGF-I receptors, hybrid receptors consisting of insulin and IGF-I receptor α-β hemireceptors can form. These are activated by IGF-I but not by insulin. The insulin receptor family of receptors differs from the erbB/Her receptors by virtue of their existence as pre-dimerized heterotetramers and their use of intermediates such as IRS and Shc proteins to couple to downstream signaling pathways.

Diabetes and Related Conditions. The epidemic of obesity occurring in the in the United States and around the world portends a significant increase in type 2 diabetes mellitus in the adult and, increasingly, in the pediatric populations. There is also growing concern regarding the prevalence of pre-diabetic conditions such as the metabolic syndrome, the incidence of which dwarfs that of clinically apparent diabetes *per se*. The hyperglycemia of type 2 diabetes results from defects in both insulin sensitivity and pancreatic β-cell function, leading to a relatively insulin-deficient state. There is also a growing appreciation that insulin resistance may play an important role in cardiac disease. A mainstay of current therapy is the use of insulin-sensitizing agents such as metformin and thiazolidinediones that act to enhance the ability of insulin to trigger appropriate cellular responses such as glucose transport in insulin target tissues. These treatments suffer, however, from a lack of mechanistic specificity, high rates of unresponsiveness (up to 30% for thiazolidinediones), and frequent side effects. Although advances are being made in the generation of islets for transplant, the time frame for the successful application of these approaches in human patients with both type 1 and type 2 disease and their ability to affect insulin resistance remains unclear. Thus, there continues to be an urgent need to design new and novel therapies to treat insulin resistance (see, *e.g.*, Alsheikh-Ali & Karas, *Amer J Cardiology*, 93:1417-8, 2004; Ovalle & Fernando, *Southern Med J.*, 95:1188-94, 2002; and Zangeneh et al., *Mayo Clinic Proc.* 78:471-479, 2003).).

The ErbB Receptor Family. The ErbB receptor family consists of four receptor tyrosine kinases: EGFR (HER-1, erbB-1), HER-2 (neu, erbB-2), HER-3 (erbB-3) and HER-4 (erbB-4). Altered expression of ErbB receptors by mutational activation, receptor overexpression, and tumor production of ligands contributes to the development and maintenance of a variety of human cancers (Olayioye et al., *Embo J.*, 19:3159-67, 2000).

The ErbB receptors are activated by several ligands with an EGF core domain (EGF-related growth factors). The exception is the HER-2 receptor, which is recruited as a preferred dimer partner with other ligand binding erbB receptors (*Id.*). The eleven mammalian EGF-like ligands are all agonists, whereas Drosophila express the ligand Argos that inhibits activation of

the EGFR (Dougall et al., *Oncogene* 9:2109-23, 1994; Hynes & Stern, *Biochim. Biophys. Acta* 1198:165-84, 1994); Tzahar & Yarden, *Biochim. Biophys. Acta* 1377:25-37, 1998).

Insulin-like growth factor 1 receptor (IGF-IR). Anti-erbB receptor antibody agents, such as the HER-2-specific antibody rhuMAb4D5 (HERCEPTIN™) have been approved for cancer therapy. Significantly, however, tumor cells may be inherently resistant, or gain resistance, to anti-erbB receptor therapies through activation of IGF-IR pathways (Chakravarti et al., *Cancer Res.* 62:200-7, 2002; Lu et al., *J. Biol. Chem.* 279:2856-65, 2004; Lu et al., *J. Natl. Cancer Inst.*, 93:1852-7, 2001). Activation of the IGF-IR by IGF-I promotes, *inter alia*, proliferation, survival, transformation, metastasis, and angiogenesis (Baserga, *Hum. Pathol.* 31, 275-6, 2000; Wang & Sun, *Curr. Cancer Drug Targets* 2:191-207, 2002), and signaling through both IGF-IR and EGF receptors is central to tumorigenesis. IGF-IR is in the same receptor family as the insulin receptor.

Herstatin. Although the HER-2 receptor does not directly bind EGF-like ligands, a secreted product of an HER-2 alternative transcript, Herstatin, binds with high affinity to the ectodomains of all members of the EGF receptor family, including EGFR/HER1/*erb*B1; HER2/*neu*/*erb*B2, HER3/erbB3, and HER4/*erb*B4, and to ΔEGFR and IGF-IR (Shamieh et al., *FEBS Letters*, 568:163-166, 2004). Herstatin was originally cloned from ovarian cancer cells, and consists of a segment (340 amino acids identical to the N-terminal subdomains I and II) of the HER-2 ectodomain, followed by 79 amino acids, encoded by intron 8 that function as a receptor binding domain (RBD) (Doherty et al., *Proc. Natl. Acad. Sci. USA* 96:10869-74, 1999). Herstatin blocks homomeric and heteromeric ErbB receptor interactions (*e.g.*, dimerization and activation), inhibits signaling by EGR ligands and by IGF-1 (*e.g.*, inhibits activation of the PI3K/Akt pathway initiated by EGF, TGFα, Heregulin and IGF-1) (Doherty et al., *Proc Natl Acad Sci.*, 96:10869-10874, 1999; Azios et al., *Oncogene*, 20:5199-5209, 2001; Justman & Clinton, *J Biol Chem.*, 277:20618-20624, 2002; Jhabvala-Romero et al., *Oncogene*, 22:8178-8186, 2003; and Shamieh et al., *supra*), causes growth arrest, and has utility as an anti-cancer agent (*Id.*, Azios et al., *Oncogene* 20:5199-209, 2001; Jhabvala-Romero et al., *Oncogene* 22:8178-86, 2003; Justman & Clinton, *J. Biol. Chem.* 277:20618-24, 2002).

There is, therefore, a need in the art to further investigate and characterize the interactions among the IR, the erbB family receptors, and the IGF-I receptor, and to identify modulators of the signaling mediated by these receptors.

There is a pronounced need in the art to identify and develop IR modulators as therapeutic agents.

There is a pronounced need in the art to design new and novel therapies to treat insulin resistance.

There is a need in the art to further assess and exploit the receptor-modulating utilities of Herstatin.

SUMMARY OF THE INVENTION

The present invention relates to therapeutic molecules and compositions for modulation of the insulin receptor (IR) and IR-mediated intracellular signaling by administration of an isoform of a cell surface receptor, and in preferred aspects, to administration of Herstatin, which is an example of such a cell surface receptor isoform. Aspects of the invention are based upon the discovery that the insulin receptor (IR) is a target of Herstatin, which specifically binds to the IR with nM affinity. According to preferred aspects of the present invention, Herstatin alters the landscape of IR-mediated signaling, exerting a positive effect on IR expression, and substantially increasing IR-mediated ERK pathway activation. The MEK (MAPK kinase)-ERK pathway has been shown to be significantly involved in glucose transport (*e.g.*, Harmon et al., *Am. J. Physiol. Endocrinol. Metab.*, 287:E758-E766, 2004).

In particular aspects, Herstatin was shown herein to bind at nM concentrations to cell-surface IR, to up-regulate basal IR expression by several-fold, and to induce the accumulation of pro-IR.

In additional aspects, and with respect to signal transduction, Herstatin was shown herein to substantially *(e.g.*, >40-fold) stimulate insulin activation of the ERK pathway, but to have little effect on insulin-stimulated activation of the PI3K/Akt pathway.

In further aspects, these changes in insulin signaling were shown herein to be accompanied by about a 4-fold *decrease* in IGF-IR expression, a decrease in the apparent serine phosphorylation state of IRS-1, and a slight decrease in IRS-2 levels as well as a decrease in apparent serine phosphorylation of IRS-2.

Therefore, according to particular aspects of the present invention, Herstatin, a cell surface receptor isoform, has substantial utility for modulating insulin signaling in cells expressing IR.

Preferred aspects of the present invention thus provide novel therapeutic methods and pharmaceutical compositions comprising a cell surface receptor isoform (*e.g.*, Herstatin, and/or variants thereof) for modulating IR, and IR-mediated signal transduction.

Alternative preferred aspects provide for a novel use of Herstatin in therapeutic methods and pharmaceutical compositions for treating various diseases associated with or characterized

by alterations in insulin sensitivity or resistance (*e.g.*, conditions or diseases characterized by altered IR expression and/or altered IR-related signaling).

In preferred embodiments, the invention provides novel methods and compositions for the treatment of conditions or diseases associated with altered IR expression or altered IR-mediated signaling, including but not limited to at least one of insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and cancer.

Additional aspects provide novel methods of targeted drug delivery.

Methods of treatment. Particularly preferred embodiments provide a method for treating or modulating a condition having an aspect related to, or associated with, or characterized by altered IR expression or altered IR-mediated signaling at a cellular level, comprising administering to a subject having such a condition, a therapeutically effective amount of a cell surface receptor isoform such as Herstatin, or a variant thereof (*e.g.*, a therapeutically effective amount of a Int8 RBD polypeptide, or a variant thereof), that binds to the extracellular domain of cellular target IR. Preferably, the condition is selected from the group consisting of insulin resistance, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, diabetes-associated lipid metabolism disorders, neurodegenerative disorders, and combinations thereof. In alternative related embodiments, the cell further expresses a target receptor selected from the group consisting of: EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4); IGF-IR and combinations thereof.

Alternative related preferred embodiments further comprise administering a therapeutically effective amount of a molecule such as a small molecule, protein, peptide or receptor-specific antibody that binds to the extracellular domain of a target receptor selected from the group consisting of: IR, EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), and IGF-IR.

Preferably, the methods further comprise administration of the cell surface receptor isoforms of this invention in combination with a therapeutically effective amount of an agent selected from the group consisting of: insulin, insulin-sensitizing agents, insulin secretogogues, and combinations thereof. Preferably, the insulin-sensitizing agent is selected from the group consisting of biguanides, metformin, thiazolidinediones (glitazones), and combinations thereof. Preferably, the insulin secretogogue is selected from the group consisting of sulfonylureas, meglitinides, and combinations thereof.

Pharmaceutical compositions. Additional preferred embodiments provide a pharmaceutical composition for treating a condition having an aspect related to, or associated with or characterized by altered IR expression or altered IR-mediated signaling at a cellular level, comprising, along with a pharmaceutically acceptable carrier or excipient, a cell surface receptor isoform such as Herstatin, or a variant thereof (*e.g.*, a Int8 RBD polypeptide, or a variant thereof), that binds to the extracellular domain of a cellular target IR. Preferably, the condition is selected from the group consisting of insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof. In alternative related preferred embodiments, the targeted cell further expresses a target receptor selected from the group consisting of: EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4); IGF-IR, and combinations thereof. Preferably, the pharmaceutical composition further comprises an agent selected from the group consisting of: insulin, insulin-sensitizing agents, insulin secretogogues, and combinations thereof. Preferably, the insulin-sensitizing agent is selected from the group consisting of biguanides, metformin, thiazolidinediones (glitazones), and combinations thereof. Preferably, the insulin secretogogue is selected from the group consisting of sulfonylureas, meglitinides, and combinations thereof..

Cell targeting. Yet further preferred embodiments provide methods and compositions for targeting a therapeutic agent to a cell expressing IR, comprising attaching the therapeutic agent to the cell surface receptor isoform, such as Herstatin, or to a variant thereof (*e.g.*, a Int8 RBD polypeptide, or a variant thereof), that binds to the extracellular domain of a cellular target IR.

In related embodiments, the targeted cell further expresses a target receptor selected from the group consisting of: EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4); IGF-IR, and combinations thereof.

Preferably, in all of the above-described preferred embodiments, the Herstatin, or variant thereof, comprises a polypeptide selected from the group consisting of SEQ ID NO:2, or a fragment of SEQ ID NO:2 of about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids are present, wherein at least one N-linked glycosylation site is present, and wherein the polypeptide binds to the extracellular domain of insulin receptor with an affinity binding constant of at least 10^8 M^{-1}. In particular aspects, the Herstatin, or variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:32-42.

Preferably, the Herstatin or variant thereof comprises SEQ ID NO:32. Preferably, the Herstatin or variant thereof consists of SEQ ID NO:32.

Preferably, the Int8 RBD polypeptide, or a variant thereof comprises a polypeptide selected from the group consisting of SEQ ID NO:1, or a fragment of SEQ ID NO:1 of about 50 to 79 contiguous residues in length, wherein the polypeptide binds to the extracellular domain of insulin receptor with an affinity binding constant of at least 10^8 M^{-1}. In particular aspects, the Int8 RBD polypeptide, or a variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:21-31. Preferably, the Int8 RBD polypeptide or variant thereof comprises SEQ ID NO:21. Preferably, the Int8 RBD polypeptide or variant thereof consists of SEQ ID NO:21.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows, according to particular aspects of the present invention as described in more detail in EXAMPLE II below, that Herstatin bound at nM concentrations to 3T3 cells over-expressing insulin receptor (IR), but not to 3T3 parental cells.

Figures 2A and 2B show, according to particular aspects as described in more detail in EXAMPLE III below, that Herstatin expression up-regulated IR expression and activation in MCF-7 cells.

Figures 3A and 3B show, according to particular aspects as described in more detail in EXAMPLE IV below, that in MCF-7 cells Herstatin expression substantially amplified insulin-stimulated ERK activation.

Figures 4A, 4B, 4C and 4D show, according to particular aspects as described in more detail in EXAMPLE V below, that Herstatin altered the expression of an array of proteins that are directly involved in insulin action.

Figure 5 shows, according to particular aspects, that the EGFR inhibitor AS1478 does not affect insulin signaling.

Figure 6 shows, according to particular aspects, that inhibition of the EGF receptor with an EGF receptor-specific inhibitor does not lead to an increase in insulin receptor.

DETAILED DESCRIPTION OF THE INVENTION

Herstatin is an example of a cell surface receptor isoform, that may also be referred to as an alternative receptor product or an intron fusion protein, which functions as a receptor ligand, and functions as a secreted ligand that inhibits members of the EGF receptor family. Herstatin binds with high affinity to all members of the EGF receptor family, including

EGFR/HER1/*erb*B1, HER2/*neu*/*erb*B2, HER3/erbB3, HER4/*erb*B4, and to ΔEGFR, and further binds to the IGF-IR.

The present invention discloses for the first time that the insulin receptor (IR) is a target of the cell surface receptor isoform, Herstatin, which specifically binds to the IR with nM affinity. According to preferred aspects of the present invention, Herstatin binds at nM concentrations to cell-surface IR, and further modulates insulin signaling in cells (*e.g.*, MCF-7 human breast cancer cells, etc) expressing IR.

Herstatin is disclosed herein to alter expression of the IR and in particular to up-regulate basal IR expression by several-fold, and induce the accumulation of pro-IR.

Herstatin is further disclosed herein to modulate insulin activation. Herstatin stimulates insulin activation of the ERK pathway in a range of about 5- to about 80-fold, while having a more modest to little effect on insulin-stimulated (IR-mediated) activation of the PI3K/Akt pathway.

Significantly, these changes in insulin signaling were shown herein to be accompanied by a *decrease* in IGF-IR expression in the range of about a 2- to about a 10-fold decrease, a decrease in the apparent serine phosphorylation state of IRS-1, and a slight decrease in IRS-2 levels as well as a decrease in apparent serine phosphorylation of IRS-2.

Therefore, preferred aspects of the present invention provide for uses of Herstatin in novel methods and compositions for treating a condition having an aspect related to, or associated with or characterized by altered IR expression or IR-mediated signal transduction.

The instant description and Examples, in various aspects, disclose the ability of Herstatin to modulate insulin action in cell models (*e.g.*, a breast cancer cell model that consists of the well-characterized MCF-7 human breast cancer cell line, and two derivative clones that express human Herstatin from a stably transfected expression vector).

In particular aspects, Herstatin binding to cell-surface IR was investigated using IR-expressing 3T3 cells (IRA-3T3). Moreover, the effects of Herstatin on the expression and activation of the IR itself, and upon the expression and activation of the major signaling pathways that emanate from the activated insulin receptor (*e.g.*, the ERK pathway and the PI3K/Akt pathway) were investigated in MCF-7 and in Herstatin-expressing MCF-7 cells. All of the individual assays were repeated a minimum of three times with similar, if not identical, results, and many of the findings have been replicated and confirmed in experiments with an independent Herstatin-expressing MCF-7 clone.

According to preferred aspects of the present invention, Herstatin upregulates IR expression and IR-mediated signal transduction (*e.g.*, substantially (>40-fold) stimulating insulin

activation of the ERK pathway). Therefore, Herstatin and/or RBD Int8 polypeptides, and Herstatin- and/or RBD Int8 polypeptide-based agents (*e.g.*, conjugates with drugs, toxins, radionuclides, etc.) have utility as therapeutic agents for treatment of diseases or conditions having an aspect related to, or associated with or characterized by altered IR expression or altered IR-mediated signaling at a cellular level (*e.g.*, insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof).

Preferred aspects provide novel methods and compositions for treating cellular insulin resistance (for discussion of insulin resistance see, *e.g.*, Alsheikh-Ali & Karas, *Amer J Cardiology*, 93:1417-8, 2004; Ovalle & Fernando, *Southern Med J.*, 95:1188-94, 2002; and Zangeneh et al., *Mayo Clinic Proc.* 78:471-479, 2003).

According to additional preferred aspects, Herstatin and/or Herstatin-based agents can be used to target IR-expressing cells and/or modulate IR-mediated signaling.

DEFINITIONS

"Herstatin," an example of a cell surface receptor isoform (also referred to as an intron fusion protein) refers to the polypeptides of SEQ ID NO:2 (including SEQ ID NOS:32-42), and additionally includes functional (*e.g.*, target receptor-binding) variants (including conservative amino acid sequence variants as described herein), fragments, muteins, derivatives and fusion proteins thereof.

As used herein, an isoform of a cell surface receptor (also referred to herein as a CSR isoform), such as an isoform of a receptor tyrosine kinase, refers to a receptor that lacks a domain or portion thereof sufficient to alter or modulate a biological activity of the receptor or modulate a biological activity compared to a wildtype and/or predominant form of the receptor. A CSR isoform refers to a receptor that lacks a domain or portion of a domain sufficient to alter or modulate a biological activity of the receptor, for example the insulin receptor. Generally, a biological activity is altered in an isoform at least 0.1, 0.5, 1, 2, 3, 4, 5, or 10-fold compared to a wildtype and/or predominant form of the receptor. Typically, a biological activity is altered 10-, 20-, 50-, 100- or 1000-fold or more. With reference to an isoform, alteration of activity refers to difference in activity between the particular isoform, which is shortened, compared to the unshortened form of the receptor. Alteration of biological activity includes an enhancement or a reduction of activity. In particular embodiments, alteration of a biological activity is a reduction in the activity. In particular embodiments, an alteration of a biological activity is a reduction in

biological activity, and the reduction can be at least 0.1 0.5 1, 2, 3, 4, 5, or 10-fold compared to a wildtype and/or predominant form of the receptor. Typically, a biological activity is reduced 5, 10, 20, 50, 100 or 1000-fold or more. Reference herein to a CSR isoform with altered activity refers to the alteration in an activity by virtue of the different structure or sequence of the CSR isoform compared to a cognate receptor.

Reference herein to modulating the activity of a target cell surface receptor means that a CSR isoform interacts in some manner with the target receptor and activity, such as ligand binding or dimerization or other signal-transduction-related activity is altered.

Intron fusion proteins (IFPs) are exemplary CSR isoforms. IFPs, for purposes herein include natural and combinatorial IFPs. A natural IFP refers to a polypeptide that is encoded by an alternatively spliced RNA that contains one or more amino acids encoded by an intron operatively linked to one or more portions of the polypeptide encoded by one or more exons of a gene. Alternatively spliced mRNA is one that is isolated or is one that can be prepared synthetically by joining splice donor and acceptor sites in a gene. A natural IFP contains one or more amino acids and/or one or more stop codons encoded by an intron sequence. A combinatorial IFP refers to a polypeptide that is shortened compared to a wildtype or predominant form of a polypeptide. Typically, the shortening removes one or more domains or a portion thereof from a polypeptide such that a biological activity is altered. Combinatorial IFPs often mimic a natural IFP in that one or more domains or a portion thereof that is/are deleted in a natural IFP derived from the same gene sequence or derived from a gene sequence in a related gene family.

As used herein, natural with reference to IFP, refers to any protein, polypeptide or peptide or fragment thereof (by virtue of the presence of the appropriate splice acceptor/donor sites) that is encoded within the genome of an animal and/or is produced or generated in an animal or that could be produced from a gene. Natural IFPs include allelic variant. IFPs can be modified post-translationally.

"RBD Int8 polypeptide" refers to the polypeptides of SEQ ID NO:1 (including SEQ ID NOS:21-31), and additionally includes functional (*e.g.*, target receptor-binding) variants (including conservative amino acid sequence variants as described herein), fragments, muteins, derivatives and fusion proteins thereof.

"Mutant RBD Int8 polypeptide" or "mutant Int8 RBD polypeptide" refers to the intron 8-encoded receptor binding domain variants (with an Arg to Ile mutation at residue 31 thereof) of SEQ ID NO:3), and additionally includes functional (*e.g.*, target receptor non-binding) variants (including conservative amino acid sequence variants as described herein), fragments, muteins,

derivatives and fusion proteins thereof. Representative, corresponding Herstatin variants (Arg to Ile mutation at residue 371) are given as SEQ ID NO:4.

"EGFR," "HER-1" or "erbB-1" refer to the art-recognized human epidermal growth factor receptor, erbB-1 (cDNA: NM_005228, SEQ ID NO:5; protein: NP_005219, SEQ ID NO:6), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

"ΔEGFR" refers to the art-recognized receptor, ΔEGFR (cDNA: SEQ ID NO:7; protein: SEQ ID NO:8) (*see* Ekstrand et al., *PNAS* 89:4309-4313, 1992; and Nishikawa et al., *PNAS* 91:7727-7731, 1994) (comprising a deletion in the ECD; cDNA positions 275 through 1075, corresponding to exons 2-7 of the EGFR gene), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

"HER-2" or "erbB-2" refers to the art-recognized human receptor, erbB-2 (cDNA: NM_004448, SEQ ID NO:9; protein: NP_004439, SEQ ID NO:10), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

"HER-3" or "erbB-3" refers to the art-recognized human receptor, erbB-3 (cDNA: NM_001982, SEQ ID NO:11; protein: NP_001973, SEQ ID NO:12), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

The phrase "mutant form of HER-3" refers to a HER-3 protein having a substitution of Glu for Gly in the ectodomain of HER-3 corresponding to a single point mutation at nucleotide position 1877 ("a" instead of "g" at this position), resulting in substitution of Glu instead of Gly at residue position 560) (cDNA: SEQ ID NO:13; protein: SEQ ID NO:14).

"HER-4" or "erbB-4" refers to the art-recognized human receptor, erbB-4 (cDNA: NM_005235, SEQ ID NO:15; protein: NP_005226, SEQ ID NO:16), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

"IGF-IR" refers to the art recognized insulin-like growth factor I receptor (cDNA: NM_000875, SEQ ID NO:17; protein: NP_000866, SEQ ID NO:18), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

"Insulin receptor" or IR refers to the art-recognized insulin receptor (cDNA: NM_000208, SEQ ID NO:19; protein: NP_000199, SEQ ID NO:20), and including Herstatin-, and/or Int8 RBD polypeptide-binding variants thereof.

TABLE 1. Summary of key SEQ ID NOS and accession numbers:

MOLECULE	cDNA	PROTEIN
RBD Int8 polypeptide(s))		SEQ ID NO:1
Herstatin(s)		SEQ ID NO:2 SEQ ID NOS:32-42
Mutant Int8 RBD polypeptide(s)		SEQ ID NO:3 SEQ ID NOS:21-31
Mutant Herstatin(s)		SEQ ID NO:4
EGFR (HER-1 or erbB-1)	SEQ ID NO:5 (NM_005228)	SEQ ID NO:6 (NP_005219)
ΔEGFR	SEQ ID NO:7	SEQ ID NO:8
HER-2 (erbB-2)	SEQ ID NO:9 (NM_004448)	SEQ ID NO:10 (NP_004439)
HER-3 (erbB-3)	SEQ ID NO:11 (NM_001982)	SEQ ID NO:12 (NP_001973)
Mutant form of HER-3	SEQ ID NO:13	SEQ ID NO:14
HER-4 (erbB-4)	SEQ ID NO:15 (NM_005235)	SEQ ID NO:16 (NP_005226)
IGF-IR	SEQ ID NO:17 (NM_000875)	SEQ ID NO:18 (NP_000866)
Insulin receptor (IR)	SEQ ID NO:19 (NM_000208)	SEQ ID NO:20 (NP_000199.1)

Cell Surface Receptor (CSR) Isoforms

Provided herein are cell surface receptor (CSR) isoforms (including intron fusion proteins; IFPs) having the novel biological activity of altering IR expression or altered IR mediated signaling. The CSR isoforms differ from the cognate receptors in that there are insertions and/or deletions, and the resulting CSR isoforms exhibit a difference in one or more activities or functions compared to the cognate receptor. Such differences include, for example elimination of all or part of a transmembrane domain, and/or a change in a biological activity of the CSR (e.g., as disclosed herein, the ability to modulate insulin receptor (IR) expression or IR-mediated signaling). The CSR isoforms provided herein can be used for modulating the activity of a cell surface receptor (e.g., the IR). They also can be used as targeting agents (*e.g.*, targeting

IR) for delivery of molecules, such as drugs or toxins or nucleic acids, to targeted cells or tissues.

A CSR isoform refers to a receptor that lacks a domain or portion of a domain sufficient to alter a biological activity (e.g., an activity with respect to the IR). Thus, an isoform may differ from a wildtype and/or predominant form of the receptor, in that it lacks one or more biological activities of the receptor. Additionally, CSR isoforms can contain a new domain and/or biological function as compared to a wildtype and/or predominant form of the receptor. For example, intron-encoded amino acids can introduce a new domain or portion thereof into a CSR isoform. Biological activities that can be altered (or gained) include, but are not limited to, protein-protein interactions such as dimerization, multimerization and complex formation, specificity and/or affinity for ligand, cellular localization and relocalization, membrane anchoring, enzymatic activity such as kinase activity, response to regulatory molecules including regulatory proteins, cofactors, and other signaling molecules, such as in a signal transduction pathway. Generally, a biological activity is altered in an isoform at least 0.1, 0.5, 1, 2, 3, 4, 5, or 10-fold as compared to a wildtype and/or predominant form of the receptor. Typically, a biological activity is altered 10, 20, 50, 100 or 1000-fold or more. For example, an isoform can be reduced with respect to a particular biological activity.

CSR isoforms can also modulate an activity of a wildtype and/or predominant form of the cognate receptor. For example, a CSR isoform can interact directly or indirectly with a CSR isoform and modulate a biological activity of the cognate receptor. Biological activities that can be altered include, but are not limited to, protein-protein interactions such as dimerization, multimerization and complex formation, specificity and/or affinity for ligand, cellular localization and relocalization, membrane anchoring, enzymatic activity such as kinase activity, response to regulatory molecules including regulatory proteins, cofactors, and other signaling molecules, such as in a signal transduction pathway.

A CSR isoform can interact directly or indirectly with a cell surface receptor to cause or participate in a biological effect, such as by modulating a biological activity of the cell surface receptor (e.g., in the instant case, the IR). A CSR isoform also can interact independently of a cell surface receptor to cause a biological effect, such as by initiating or inhibiting a signal transduction pathway. For example, a CSR isoform can initiate a signal transduction pathway and enhance or promote cellular metabolism. In another example, a CSR isoform can interact with the cell surface receptor as a ligand, causing a biological effect for example by inhibiting a signal transduction pathway that can promote or alter a cellular response to insulin. Hence, the isoforms provided herein can function as cell surface receptor ligands in that they interact with

the targeted receptor in the same manner that a cognate ligand interacts with and alters receptor activity. The isoforms can bind as a ligand, but not necessarily to the ligand binding site, and can serve to block receptor dimerization. They act as ligands in the sense that they interact with the receptor. The CSR isoforms also can act by binding to ligands for the receptor and/or by preventing receptor activities, such as dimerization.

For example, a CSR isoform can compete with a CSR for ligand binding. A CSR isoform can act as a dominant negative inhibitor, for example, when complexed with a CSR. A CSR isoform can act as a dominant negative inhibitor or as a competitive inhibitor of a CSR, for example, by complexing with a CSR isoform and altering the ability of the CSR to multimerize (*e.g,* dimerize or trimerize) with other CSRs. A CSR isoform can compete with a CSR for interactions with other polypeptides and cofactors in a signal transduction pathway.

The cell surface isoforms and families of isoforms provided herein include, for example, isoforms of the HER-2 receptor (*e.g.*, Herstatin), IR, etc. Pharmaceutical compositions containing one or more different CSR isoforms are provided. Also provided are methods of treatment of diseases and conditions by administering the pharmaceutical compositions or delivering a CSR isoform, such by administering the isoform protein (polypeptide, etc), and/or by administration of a vector that encodes the isoform. Administration, by either means, can be effected *in vivo* or *ex vivo*. Also provided are methods for expressing, isolating and formulating CSR isoforms.

Herstatin and/or RBD Int8 polypeptides and therapeutic agents

In preferred aspects, the present invention provides for Herstatin (*e.g.*, the sequences of SEQ ID NO:2) and polypeptides thereof that bind to a *insulin receptor* (*IR*) as a target receptor (specifically, or in addition to the known targets: EGFR, HER-2, HER-3, DEGFR, HER-4 and IGF-IR). Also provided are RBD Int8 polypeptides (*e.g.*, the sequences of SEQ ID NO:1) and receptor-binding polypeptides thereof that bind to a *insulin receptor* as a target receptor (specifically, or in addition to the known targets EGFR, HER-2, HER-3, DEGFR, HER-4 and IGF-IR).

Preferably, the Herstatin and/or RBD Int8 polypeptides comprise an amino acid sequence of SEQ ID NO:1 (or of SEQ ID NO:1 having from 1, to about 3, to about 5, to about 10, or to about 20 conservative amino acid substitutions), or a fragment of a sequence of SEQ ID NO:1 (or a fragment of SEQ ID NO:1 having from 1, to about 3, to about 5, to about 10, or to about 20 conservative amino acid substitutions) of about 50 to 79 contiguous residues in length, wherein the polypeptide binds to the extracellular domain (ECD) of a target receptor (*e.g.*,

EGFR, HER-2, HER-3, DEGFR, HER-4, IGF-IR and IR (as disclosed herein)) with an affinity binding constant of at least 10^7 M^{-1}, at least 5 x 10^7 M^{-1}, or at least 10^8 M^{-1}. Preferably, the Herstatin and/or RBD Int8 polypeptide is from about 69 to 79 contiguous residues in length, with a IR affinity binding constant of at least 10^7 M^{-1}, at least 5 x 10^7 M^{-1}, or at least 10^8 M^{-1} (similar to the respective binding constants associated with the known EGFR, HER-2, HER-3, DEGFR, HER-4 and IGF-IR target receptors). Preferably, Herstatin and/or RBD Int8 polypeptide comprises a sequence of SEQ ID NO:1, or a conservative amino acid substitution variant thereof. In particular aspects, the Int8 RBD polypeptide, or a variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:21-31. Preferably, the Int8 RBD polypeptide or variant thereof comprises SEQ ID NO:21. Preferably, the Int8 RBD polypeptide or variant thereof consists of SEQ ID NO:21.

Preferably, the Herstatin and/or RBD Int8 polypeptides comprise an amino acid sequence of SEQ ID NO:2 (or of SEQ ID NO:2 having from 1, to about 3, to about 5, to about 10, or to about 20 conservative amino acid substitutions), or a fragment of a sequence of SEQ ID NO:2 (or a fragment of SEQ ID NO:2 having from 1, to about 3, to about 5, to about 10, or to about 20 conservative amino acid substitutions) of about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids are present, and wherein the polypeptide binds to the extracellular domain (ECD) of a IR with an affinity binding constant of at least 10^7 M^{-1}, at least 5 x 10^7 M^{-1}, or at least 10^8 M^{-1} (similar to the respective binding constants associated with the known EGFR, HER-2, HER-3, DEGFR, HER-4 and IGF-IR target receptors). Preferably, the Herstatin and/or RBD Int8 polypeptide is from about 350 to 419 contiguous residues in length, wherein the polypeptide binds to the extracellular domain (ECD) of a IR with an affinity binding constant of at least 10^7 M^{-1}, at least 5 x 10^7 M^{-1}, or at least 10^8 M^{-1} (similar to the respective binding constants associated with the known EGFR, HER-2, HER-3, DEGFR, HER-4 and IGF-IR target receptors). Preferably, comprises a sequence of SEQ ID NO:2, or a conservative amino acid substitution variant thereof. In particular aspects, the Herstatin, or variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:32-42. Preferably, the Herstatin or variant thereof comprises SEQ ID NO:32. Preferably, the Herstatin or variant thereof consists of SEQ ID NO:32.

Biologically Active Variants

Variants of Herstatin and/or RBD Int8 polypeptide have substantial utility in various aspects of the present invention. Variants can be naturally or non-naturally occurring. Naturally occurring variants are found in humans or other species and comprise amino acid sequences

which are substantially identical to the amino acid sequences shown in SEQ ID NO:1 or SEQ ID NO:2, and include natural sequence polymorphisms. Species homologs of the protein can be obtained using subgenomic polynucleotides of the invention, as described below, to make suitable probes or primers for screening cDNA expression libraries from other species, such as mice, monkeys, yeast, or bacteria, identifying cDNAs which encode homologs of the protein, and expressing the cDNAs as is known in the art.

Non-naturally occurring variants which retain substantially the same biological activities as naturally occurring protein variants, including the target RBD activity and the modulation of target receptor signaling activity, are also included here. Preferably, naturally or non-naturally occurring variants have amino acid sequences which are at least 85%, 90%, or 95% identical to the amino acid sequence shown in SEQ ID NOS:1 or 2. More preferably, the molecules are at least 98% or 99% identical. Percent identity is determined using any method known in the art. A non-limiting example is the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 1. The Smith-Waterman homology search algorithm is taught in Smith and Waterman, *Adv. Appl. Math.* 2:482-489, 1981.

As used herein, "amino acid residue" refers to an amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are generally in the "L" isomeric form. Residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH_2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxyl terminus of a polypeptide. In keeping with standard polypeptide nomenclature described in *J. Biol. Chem.*, 243:3552-59 (1969) and adopted at 37 C.F.R. §§ 1.821-1.822, abbreviations for amino acid residues are shown in Table 1:

TABLE 1 – Table of Correspondence

SYMBOL		
1-Letter	**3-Letter**	**AMINO ACID**
Y	Tyr	Tyrosine
G	Gly	Glycine
F	Phe	Phenylalanine
M	Met	Methionine

SYMBOL		
A	Ala	Alanine
S	Ser	Serine
I	Ile	Isoleucine
L	Leu	Leucine
T	Thr	Threonine
V	Val	Valine
P	Pro	Praline
K	Lys	Lysine
H	His	Histidine
Q	Gln	Glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	Tryptophan
R	Arg	Arginine
D	Asp	aspartic acid
N	Asn	Asparagines
B	Asx	Asn and/or Asp
C	Cys	Cysteine
X	Xaa	Unknown or other

It should be noted that all amino acid residue sequences represented herein by a formula have a left to right orientation in the conventional direction of amino-terminus to carboxyl-terminus. In addition, the phrase "amino acid residue" is defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those referred to in 37 C.F.R. §§ 1.821-1.822, and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino acid residues or to an amino-terminal group such as NH_2 or to a carboxyl-terminal group such as COOH.

Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological or immunological activity can be found using computer programs well known in the art, such as DNASTAR™ software. Preferably, amino acid changes in the protein variants disclosed herein are conservative amino acid changes, *i.e.*, substitutions of similarly charged or uncharged amino acids. A conservative amino acid change

involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cystine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids.

In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and generally can be made without altering a biological activity of a resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p.224).

Such substitutions may be made in accordance with those set forth in TABLE 2 as follows:

TABLE 2

Original residue	**Conservative substitution**
Ala (A)	Gly; Ser
Arg (R)	Lys
Asn (N)	Gln; His
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
Gly (G)	Ala; Pro
His (H)	Asn; Gln
Ile (I)	Leu; Val
Leu (L)	Ile; Val
Lys (K)	Arg; Gln; Glu
Met (M)	Leu; Tyr; Ile
Phe (F)	Met; Leu; Tyr
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp; Phe
Val (V)	Ile; Leu

Other substitutions also are permissible and can be determined empirically or in accord with other known conservative (or non-conservative) substitutions.

Variants of the Herstatin and/or RBD Int8 polypeptide disclosed herein include glycosylated forms, aggregative conjugates with other molecules, and covalent conjugates with unrelated chemical moieties (*e.g.*, pegylated molecules). Covalent variants can be prepared by linking functionalities to groups which are found in the amino acid chain or at the N- or C-terminal residue, as is known in the art. Variants also include allelic variants, species variants, and muteins. Truncations or deletions of regions which do not affect functional activity of the proteins are also variants.

A subset of mutants, called muteins, is a group of polypeptides in which neutral amino acids, such as serines, are substituted for cysteine residues which do not participate in disulfide bonds. These mutants may be stable over a broader temperature range than native secreted proteins (Mark *et al.*, United States Patent 4,959,314).

Preferably, amino acid changes in the Herstatin and/or RBD Int8 polypeptide variants are conservative amino acid changes, *i.e.*, substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cystine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids.

It is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the biological properties of the resulting secreted protein or polypeptide variant. Properties and functions of Herstatin and/or RBD Int8 polypeptide protein or polypeptide variants are of the same type as a protein comprising the amino acid sequence encoded by the nucleotide sequences shown in SEQ ID NO:1 or 2, although the properties and functions of variants can differ in degree.

Herstatin and/or RBD Int8 polypeptide variants include glycosylated forms, aggregative conjugates with other molecules, and covalent conjugates with unrelated chemical moieties (*e.g.*, pegylated molecules). Herstatin and/or RBD Int8 polypeptide variants also include allelic variants (*e.g.*, polymorphisms), species variants, and muteins. Truncations or deletions of

regions which do not preclude functional activity of the proteins are also variants. Covalent variants can be prepared by linking functionalities to groups which are found in the amino acid chain or at the N- or C-terminal residue, as is known in the art.

It will be recognized in the art that some amino acid sequence of the Herstatin and/or RBD Int8 polypeptides of the invention can be varied without significant effect on the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there are critical areas on the protein which determine activity. In general, it is possible to replace residues that form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the protein. The replacement of amino acids can also change the selectivity of binding to cell surface receptors (Ostade et al., *Nature 361*:266-268, 1993). Thus, the Herstatin and/or RBD Int8 polypeptides of the present invention may include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation.

Of particular interest are substitutions of charged amino acids with another charged amino acid and with neutral or negatively charged amino acids. The latter results in proteins with reduced positive charge to improve the characteristics of the disclosed protein. The prevention of aggregation is highly desirable. Aggregation of proteins not only results in a loss of activity but can also be problematic when preparing pharmaceutical formulations, because they can be immunogenic (Pinckard et al., *Clin. Exp. Immunol.* 2:331-340, 1967; Robbins et al., *Diabetes* 36:838-845, 1987; Cleland et al., *Crit. Rev. Therapeutic Drug Carrier Systems* 10:307-377, 1993).

Amino acids in the Herstatin and/or RBD Int8 polypeptides of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science 244*:1081-1085, 1989). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as binding to a natural or synthetic binding partner. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith et al., *J. Mol. Biol. 224*:899-904, 1992 and de Vos et al. *Science 255*:306-312,1992).

As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein. Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors,

including those described above. Generally speaking, the number of substitutions for any given Herstatin and/or RBD Int8 polypeptide will not be more than 50, 40, 30, 25, 20, 15, 10, 5 or 3.

In addition, pegylation of Herstatin and/or RBD Int8 polypeptides and/or muteins is expected to provide such improved properties as increased half-life, solubility, and protease resistance. Pegylation is well known in the art.

Fusion Proteins

Fusion proteins comprising proteins or polypeptide fragments of Herstatin and/or RBD Int8 polypeptide can also be constructed. Fusion proteins are useful for generating antibodies against amino acid sequences and for use in various targeting and assay systems. For example, fusion proteins can be used to identify proteins which interact with a Herstatin and/or RBD Int8 polypeptide of the invention or which interfere with its biological function. Physical methods, such as protein affinity chromatography, or library-based assays for protein-protein interactions, such as the yeast two-hybrid or phage display systems, can also be used for this purpose. Such methods are well known in the art and can also be used as drug screens. Fusion proteins comprising a signal sequence can be used.

A fusion protein comprises two protein segments fused together by means of a peptide bond. Amino acid sequences for use in fusion proteins of the invention can be utilize the amino acid sequence shown in SEQ ID NOS:1 or 2 or can be prepared from biologically active variants of SEQ ID NOS:1 or 2, such as those described above. The first protein segment can include of a full-length Herstatin and/or RBD Int8 polypeptide.

Other first protein segments can consist of about 50 to about 79 contiguous amino acids from SEQ ID NO:1, or, with respect to SEQ ID NO:2, from about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids of SEQ ID NO:2 are present, or from about 350 to 419 contiguous residues in length wherein the C-terminal 79 contiguous amino acids of SEQ ID NO:2 are present.

The second protein segment can be a full-length protein or a polypeptide fragment. Proteins commonly used in fusion protein construction include β-galactosidase, β-glucuronidase, green fluorescent protein (GFP), autofluorescent proteins, including blue fluorescent protein (BFP), glutathione-S-transferase (GST), luciferase, horseradish peroxidase (HRP), and chloramphenicol acetyltransferase (CAT). Additionally, epitope tags can be used in fusion protein constructions, including histidine (His) tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Other fusion constructions can

include maltose binding protein (MBP), S-tag, Lex a DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions.

These fusions can be made, for example, by covalently linking two protein segments or by standard procedures in the art of molecular biology. Recombinant DNA methods can be used to prepare fusion proteins, for example, by making a DNA construct which comprises a coding region for the protein sequence of SEQ ID NOS:1 or 2 in proper reading frame with a nucleotide encoding the second protein segment and expressing the DNA construct in a host cell, as is known in the art. Many kits for constructing fusion proteins are available from companies that supply research labs with tools for experiments, including, for example, Promega Corporation (Madison, WI), Stratagene (La Jolla, CA), Clontech (Mountain View, CA), Santa Cruz Biotechnology (Santa Cruz, CA), MBL International Corporation (MIC; Watertown, MA), and Quantum Biotechnologies (Montreal, Canada; 1-888-DNA-KITS).

Cell Targeting

According to additional preferred aspects of the present invention, cell surface receptor isoforms such as Herstatin- and/or RBD Int8 polypeptide-based agents can be used to target insulin receptor (IR) on cells (*e.g.*, insulin-resistant cells, IR-expressing cells involved with some aspect of glucose regulation or metabolism, cancer cells, etc.). Herstatin- and/or RBD Int8 polypeptide-based agents can be used to deliver a locally acting biological agent that will affect the targeted cell.

IR, in the context of the inventive targeting, is expressed on the surface of cells and is accessible (specifically, or in addition to at least one of the other known Herstatin targets: EGFR; HER-2; HER-3; HER-4, ΔEGFR and IGF-IR) to exogenous molecules. For example, where IR is present at higher levels on particular IR-bearing cells (*e.g.*, adipocytes, hepatocytes, skeletal muscle cells, pancreatic beta cells, brain/nerve cells, etc) as compared to other cells, they can be utilized as preferential targets for systemic Herstatin- and/or RBD Int8 polypeptide-based agents and therapies. The differential expression of the target receptor (*e.g.*, IR) enables the specificity of Herstatin- and/or RBD Int8 polypeptide-based agents-based therapy. Herstatin- and/or RBD Int8 polypeptide-based agents (*e.g.*, drugs, cytoxic agents, labeling agents, etc.) directed against the target receptor preferentially affect the targeted cell over normal tissue. For example, a Herstatin- or RBD Int8 polypeptide-drug conjugate that binds a IR present predominantly on particular cells (*e.g.*, adipocytes, hepatocytes, skeletal muscle cells, pancreatic beta cells, brain/nerve cells, etc) would be expected to selectively affect those cells within a treated individual. Preferably, the target receptor is accessible to the Herstatin- and/or

RBD Int8 polypeptide-based agent, and is found in substantially greater concentrations on the targeted cells (*e.g.*, adipocytes, hepatocytes, skeletal muscle cells, pancreatic beta cells, brain/nerve cells, etc) relative to other cells that don't express IR or that express IR at relatively low levels.

Therefore, the present invention includes Herstatin- and/or RBD Int8 polypeptide-based agents specific to one or more of the target receptors (*e.g.*, IR) that will enable or facilitate therapeutic treatments relating to, for example, adipocytes, hepatocytes, skeletal muscle cells, pancreatic beta cells, brain cells, etc.

In particular aspects, Herstatin- and/or RBD Int8 polypeptides are conjugated or coupled to drugs, or to toxins.

In alternate embodiments, Herstatin- and/or RBD Int8 polypeptides are conjugated or coupled to radionuclides.

Additional embodiments provide for Herstatin- and/or RBD Int8 polypeptide-coated liposomes that contain one or more biologically active compounds.

In preferred embodiments, Herstatin-mediated targeting is used to deliver drugs or other agents to adipocytes, hepatocytes, skeletal muscle cells, pancreatic beta cells, brain cells, and combinations thereof.

In alternate aspects, targeted binding of an Herstatin- and/or RBD Int8 polypeptide-agent to a cell is sufficient to modulate IR-mediated signaling, inhibit or alter growth (*e.g.*, cytostatic effects) or even kill the target cell (cytotoxic effects) if so desired. The mechanism of these activities may vary, but may involve Herstatin- and/or RBD Int8 polypeptide-dependent receptor activation, changes in receptor expression, cell-mediated cytotoxicity, activation of apoptosis, inhibition of ligand-receptor function, or provide a signal for complement fixation. In fact, Herstatin- and/or RBD Int8 polypeptide-agents may exhibit one or several such activities. In particular aspects, Herstatin- and/or RBD Int8 polypeptide-agents are cytostatic, but not cytotoxic. In particular embodiments, Herstatin- and/or RBD Int8 polypeptide-agents bind to target receptors (*e.g.*, IR, EGFR (HER-1, erbB-1); HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), ΔEGFR or IGF-IR), and modulate signaling and cellular metabolism, or are either cytoxic or cytostatic, etc.

In additional embodiments, Herstatin- and/or RBD Int8 polypeptide-agents are conjugated or coupled to a diverse array of compounds which include, but are not limited to proteins, drugs, toxins or cytotoxic agents, cytostatic agents, radionuclides, apoptotic factors (Wuest et al. 2002), anti-angiogenic compounds or other biologically active compounds which will affect cellular signaling or metabolism, inhibit the growth of or even kill the target cell or

tissue. For example, cytotoxic or cytostatic agents include, but are not limited to, diphtheria toxin and Pseudomonas exotoxin (Kreitman 2001 a; Kreitman 2001 b), ricin (Kreitman 2001 a), gelonin, doxorubicin (Ajani et al. 2000) and its derivatives, iodine-131, yttrium-90 (Witzig 2001), indium-111 (Witzig 2001), RNase (Newton and Ryback 2001), calicheamicin (Bernstein 2000), apoptotic agents, and antiangiogenic agents (Frankel et al. 2000; Brinkmann et al. 2001; Garnett 2001). According to particular aspects of the present invention, Herstatin- and/or RBD Int8 polypeptides coupled to these compounds are used to adversely affect cells displaying one or more target receptors (*e.g.*, IR, EGFR (HER-1, erbB-1); HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), ΔEGFR or IGF-IR).

Toxins can also be targeted to specific cells by incorporation of the toxin into Herstatin- and/or RBD Int8 polypeptide-coated liposomes. The Herstatin- and/or RBD Int8 polypeptide-based agent directs the liposome to the target cell where the bioactive compound is released. For example, cytotoxins in Herstatin- and/or RBD Int8 polypeptide-coated liposomes are used to treat cancer. In alternate embodiments, these targeted liposomes are loaded with DNA encoding bioactive polypeptides (*e.g.*, inducible nitric oxide synthase; Khare et al. 2001).

Prodrugs or enzymes can also be delivered to targeted cells by specific Herstatin- and/or RBD Int8 polypeptide-agents. In this case the Herstatin conjugate consists of a Herstatin- and/or RBD Int8 polypeptide-based agent coupled to a drug that can be activated once the polypeptide agent binds the target cell. Examples of this strategy using antibodies have been reviewed (Denny 2001; Xu and McLeod 2001).

Therefore, in particular embodiments, Herstatin- and/or RBD Int8 polypeptide-prodrug/enzyme conjugates targeted to one or more target receptors (*e.g.*, IR, EGFR (HER-1, erbB-1); HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), ΔEGFR or IGF-IR) have utility for the treatment of, for example, cancer and other treatable conditions discussed herein.

The specificity and high affinity of the Herstatin- and/or RBD Int8 polypeptide-based agents makes them ideal candidates for delivery of toxic agents to a specific subset of cellular targets. Preferably, one or more target receptors (*e.g.*, IR, EGFR (HER-1, erbB-1); HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), ΔEGFR or IGF-IR) are present at higher levels on the target cells (*e.g.*, cancer, tumor cells) than on non-cancer cells.

As used herein, a composition refers to any mixture. It can be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

As used herein, a combination refers to any association between or among two or more items. The combination can be two or more separate items, such as two compositions or two

collections, can be a mixture thereof, such as a single mixture of the two or more items, or any variation thereof.

As used herein, a pharmaceutical effect refers to an effect observed upon administration of an agent intended for treatment of a disease or disorder or for amelioration of the symptoms thereof.

As used herein, treatment means any manner in which the symptoms of a condition, disorder or disease or other indication, are ameliorated or otherwise beneficially altered.

As used herein therapeutic effect means an effect resulting from treatment of a subject that alters, typically improves or ameliorates the symptoms of a disease or condition or that cures a disease or condition. A therapeutically effective amount refers to the amount of a composition, molecule or compound which results in a therapeutic effect following administration to a subject.

In particular aspects, a therapeutic effect may also encompass prophylaxis of symptoms of a condition.

As used herein, the term "subject" refers to animals, including mammals, such as human beings. As used herein, a patient refers to a human subject.

As used herein, the phrase "associated with" or "characterized by" refers to certain biological aspects such as expression of a receptor or signaling by a receptor that occurs in the context of a disease or condition. Such biological aspects may or may not be causative or integral to the disease or condition but merely an aspect of the disease or condition.

As used herein, a biological activity refers to a function of a polypeptide including but not limited to complexation, dimerization, multimerization, receptor-associated kinase activity, receptor-associated protease activity, phosphorylation, dephosphorylation, autophosphorylation, ability to form complexes with other molecules, ligand binding, catalytic or enzymatic activity, activation including auto-activation and activation of other polypeptides, inhibition or modulation of another molecule's function, stimulation or inhibition of signal transduction and/or cellular responses such as cell proliferation, migration, differentiation, and growth, degradation, membrane localization, membrane binding, and oncogenesis. A biological activity can be assessed by assays described herein and by any suitable assays known to those of skill in the art, including, but not limited to *in vitro* assays, including cell-based assays, *in vivo* assays, including assays in animal models for particular diseases.

Pharmaceutical Compositions and Therapeutic Uses

Pharmaceutical compositions of the invention comprise a cell surface receptor isoform such as Herstatin and/or RBD Int8 polypeptides, or Herstatin- and/or RBD Int8 polypeptide-based agents of the claimed invention in a therapeutically effective amount. The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation is determined by routine experimentation and is within the judgment of the clinician. For purposes of the present invention, an effective dose will generally be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the Herstatin and/or RBD Int8 polypeptide constructs in the individual to which it is administered. A non-limiting example of a pharmaceutical composition is a composition that either enhances or diminishes signaling mediated by the inventive target receptors (*e.g.*, IR, EGFR, HER-2, HER-3, ΔEGFR, HER-4 and IGF-IR). Where such signaling modulates a disease-related process, modulation of the signaling would be the goal of the therapy.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which can be administered without undue toxicity. Suitable carriers can be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Pharmaceutically acceptable carriers in therapeutic compositions can include liquids such as water, saline, glycerol and ethanol. Auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, can also be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier. Pharmaceutically acceptable salts can also be present in the pharmaceutical composition, *e.g.*, mineral acid salts such as hydrochlorides, hydrobromides,

phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in *Remington's Pharmaceutical Sciences* (Mack Pub. Co., New Jersey, 1991).

Delivery Methods

Once formulated, the compositions of the invention can be administered (as proteins/polypeptides, or in the context of expression vectors for gene therapy) directly to the subject or delivered *ex vivo*, to cells derived from the subject (*e.g.*, as in *ex vivo* gene therapy). Direct delivery of the compositions will generally be accomplished by parenteral injection, *e.g.*, subcutaneously, intraperitoneally, intravenously or intramuscularly, myocardial, intratumoral, peritumoral, or to the interstitial space of a tissue. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment can be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in, for example, International Publication No. WO 93/14778. Examples of cells useful in *ex vivo* applications include, for example, stem cells, particularly hematopoetic, lymph cells, macrophages, dendritic cells, or tumor cells. Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, direct microinjection of the DNA into nuclei, and viral-mediated, such as adenovirus (and adeno-associated virus) or alphavirus, all well known in the art.

In a preferred embodiment, certain disorders (e.g., of proliferation, such as cancer, etc), can be amenable to treatment by administration of a therapeutic agent based on the provided polynucleotide or corresponding polypeptide. The therapeutic agent can be administered in conjunction with one or more other agents including, but not limited to, receptor-specific antibodies and/or other agents (*e.g.*, insulin-sensitizing agents, chemotherapeutic agents, etc). Administered "in conjunction" includes administration at the same time, or within 1 day, 12 hours, 6 hours, one hour, or less than one hour, as the other therapeutic agent(s). The compositions may be mixed for co-administration, or may be administered separately by the same or different routes.

The dose and the means of administration of the inventive pharmaceutical compositions are determined based on the specific qualities of the therapeutic composition, the condition, age, and weight of the patient, the progression of the disease, and other relevant factors. For example, administration of polynucleotide therapeutic compositions agents of the invention includes local or systemic administration, including injection, oral administration, particle gun or catheterized administration, and topical administration. The therapeutic polynucleotide composition can contain an expression construct comprising a promoter operably linked to a polynucleotide encoding, for example, about 80 to 419 (or about 350 to 419) contiguous amino acids of SEQ ID NO:2. Various methods can be used to administer the therapeutic composition directly to a specific site in the body. For example, an abnormal tissue, or small metastatic lesion is located and the therapeutic composition injected several times in several different locations within the body of the tissue, or tumor. Alternatively, arteries which serve a tissue or tumor are identified, and the therapeutic composition injected into such an artery, in order to deliver the composition directly into the tumor. A tissue or tumor that has a necrotic center is aspirated and the composition injected directly into the now empty center of the tissue or tumor. X-ray imaging is used to assist in certain of the above delivery methods.

Herstatin and/or RBD Int8 polypeptide-mediated targeted delivery of therapeutic agents to specific tissues can also be used. Receptor-mediated DNA delivery techniques are described in, for example, Findeis et al., *Trends Biotechnol.* (1993) *11*:202; Chiou et al., *Gene Therapeutics: Methods And Applications Of Direct Gene Transfer* (J.A. Wolff, ed.) (1994); Wu et al., *J. Biol. Chem.* (1988) *263*:621; Wu et al., *J. Biol. Chem.* (1994) *269*:542; Zenke et al., *Proc. Natl. Acad. Sci.* (*USA*) (1990) *87*:3655; Wu et al., *J. Biol. Chem.* (1991) *266*:338.

For gene therapy, therapeutic compositions containing a polynucleotide are administered in a range of about 100 ng to about 200 mg of DNA for local administration in a gene therapy protocol. Concentration ranges of about 500 ng to about 50 mg, about 1 mg to about 2 mg, about 5 mg to about 500 mg, and about 20 mg to about 100 mg of DNA can also be used during a gene therapy protocol. Factors such as method of action (*e.g.*, for enhancing or inhibiting levels of the encoded gene product) and efficacy of transformation and expression are considerations which will affect the dosage required for ultimate efficacy of the subgenomic polynucleotides. Where greater expression is desired over a larger area of tissue, larger amounts of subgenomic polynucleotides or the same amounts re-administered in a successive protocol of administrations, or several administrations to different adjacent or close tissue portions of, for example, a tumor site, may be required to affect a positive therapeutic outcome. In all cases,

routine experimentation in clinical trials will determine specific ranges for optimal therapeutic effect.

The therapeutic polynucleotides and polypeptides of the present invention can be delivered using gene delivery vehicles. The gene delivery vehicle can be of viral or non-viral origin (see generally, Jolly, *Cancer Gene Therapy* (1994) *1*:51; Kimura, *Human Gene Therapy* (1994) *5*:845; Connelly, *Human Gene Therapy* (1995) *1*:185; and Kaplitt, *Nature Genetics* (1994) *6*:148). Expression of such coding sequences can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence can be either constitutive or regulated.

Viral-based vectors for delivery of a desired polynucleotide and expression in a desired cell are well known in the art. Exemplary viral-based vehicles include, but are not limited to, recombinant retroviruses (see, *e.g.*, WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5, 219,740; WO 93/11230; WO 93/10218; U.S. Patent No. 4,777,127; GB Patent No. 2,200,651; EP 0 345 242; and WO 91/02805), alphavirus-based vectors (*e.g.*, Sindbis virus vectors, Semliki forest virus (ATCC VR-67; ATCC VR-1247), Ross River virus (ATCC VR-373; ATCC VR-1246) and Venezuelan equine encephalitis virus (ATCC VR-923; ATCC VR-1250; ATCC VR 1249; ATCC VR-532), and adeno-associated virus (AAV) vectors (see, *e.g.*, WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 and WO 95/00655). Administration of DNA linked to killed adenovirus as described in Curiel, *Hum. Gene Ther.* (1992) *3*:147 can also be employed.

Non-viral delivery vehicles and methods can also be employed, including, but not limited to, polycationic condensed DNA linked or unlinked to killed adenovirus alone (see, *e.g.*, Curiel, *Hum. Gene Ther.* (1992) *3*:147); ligand-linked DNA (see, *e.g.*, Wu, *J. Biol. Chem. 264*:16985 (1989)); eukaryotic cell delivery vehicles cells (see, *e.g.*, U.S. Patent No. 5,814,482; WO 95/07994; WO 96/17072; WO 95/30763; and WO 97/42338) and nucleic charge neutralization or fusion with cell membranes. Naked DNA can also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and U.S. Patent No. 5,580,859. Liposomes that can act as gene delivery vehicles are described in U.S. Patent No. 5,422,120; WO 95/13796; WO 94/23697; WO 91/14445; and EP 0524968. Additional approaches are described in Philip, *Mol. Cell Biol. 14*:2411 (1994), and in Woffendin, *Proc. Natl. Acad. Sci.* (1994) *91*:11581-11585.

Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin et al., *Proc. Natl. Acad. Sci. USA 91*(24):11581 (1994). Moreover, the coding sequence and the product of expression of such can be delivered through

deposition of photopolymerized hydrogel materials or use of ionizing radiation (see, *e.g.*, U.S. Patent No. 5,206,152 and WO 92/11033). Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun (see, *e.g.*, U.S. Patent No. 5,149,655); use of ionizing radiation for activating transferred gene (see, *e.g.*, U.S. Patent No. 5,206,152 and WO 92/11033).

Conditions Treatable

Particular aspects of the present invention, for the first time, disclose that Herstatin or Int8 RBD polypeptides, and variants thereof, can not only modulate the expression/level of cellular insulin receptors (IR) (both pro-IR and IR), but also modulate IR-mediated signal transduction (*e.g.*, ERK pathway). According to particular aspects, Herstatin or Int8 RBD polypeptides, and variants thereof can be used in therapeutic methods and pharmaceutical compositions to treat a variety of conditions having an aspect related to, or associated with altered IR expression or altered IR-mediated signaling at a cellular level. Such methods comprising administering to a subject having such a condition, a therapeutically effective amount of a Herstatin or Int8 RBD polypeptide, or a variant thereof, that binds to the extracellular domain of cellular target insulin receptor. Such methods also encompass gene delivery-related methods.

IR is well known in the art to be involved with, *inter alia*, glycemic control (*e.g.*, hyper- and hypo-glycemia) and glucose metabolism. Accordingly, conditions having an aspect related to, or associated with altered glycemic control and/or glucose metabolism are within the scope of treatable conditions according to the present invention. Such conditions include, but are not limited to insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof.

Insulin resistance syndrome has become the major health problem of our times, and is associated with obesity, dyslipidemia, atherosclerosis, hypertension, and type-2 diabetes shorten life spans, and hyperandrogenism with polycystic ovarian syndrome affect quality of life and fertility in increasing numbers of women (see, *e.g.*, Ten & Maclaren, *J. Clin Endocrinol Metab.*, 89:2526-2539, 2004; and see Le Roith 7 Zick, *Diabetes Care* 24:588-597, 2001; both incorporated herein by reference). In particular preferred aspects, Herstatin or Int8 RBD polypeptide, or variants thereof can be used to treat insulin resistance syndrome.

Insulin resistance and associated abnormalities are believed to have a role in pregnancy induced hypertension (new-onset hypertension), and many features of the insulin resistance syndrome are associated with this condition (see, *e.g.*, Seely & Solomon, *J. Clin. Endocrinol. Metab.*, 88:2393-2398, 2003; incorporated herein by reference). According to the present invention, Herstatin or Int8 RBD polypeptide, or variants thereof can be used to treat hypertension and new-onset hypertension.

In prolonged critical illness neuroendocrine changes lead to more extensive metabolic changes. For example, insulin resistance and hyperglycemia are associated with critical illness (*e.g.*, in surgically critically ill populations with or without diabetes, post-myocardial infarction in patients with diabetes, etc.) (see, *e.g.*, Ronbinson & H. van Soeren, AACN Clinical Issues, 15:45-62, 2004; incorporated herein by reference). According to the present invention, Herstatin or Int8 RBD polypeptide, or variants thereof can be used to treat critical illness.

Significantly, impairment of insulin signaling in the brain has been linked, on the basis of studies using IR-knockout (NIRKO) mice, to neurodegenerative diseases. NIRKO mice exhibit a complete loss of insulin-mediated activation of phosphatidylinositol 3-kinase and insulin-mediated inhibition of neuronal apoptosis, resulting in markedly reduced phosphorylation of Akt and GSK3 β and leading to a substantially increased phosphorylation of the microtubule-associated protein Tau, a hallmark of neurodegenerative diseases (*e.g.*, Alzheimer's disease) (see, *e.g.*, Schubert et al., PNAS 101:3100-3105, 2004, incorporated herein by reference). According to the present invention, Herstatin or Int8 RBD polypeptide, or variants thereof can be used to treat to neurodegenerative diseases (*e.g.*, Alzheimer's disease).

Combination Therapies

According to additional preferred aspects of the invention, Herstatin-related treatment of conditions having an aspect related to, or characterized by altered glycemic control and/or glucose metabolism, including, but not limited to insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, and combinations thereof, may further comprise administration of another therapeutic agent.

For example, the inventive treatment methods may further comprise administering a therapeutically effective amount of a receptor-specific antibody that binds to the extracellular domain of a target receptor selected from the group consisting of: IR, EGFR (HER-1, erbB-1); □EGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), and IGF-IR.

Alternatively, the inventive treatment methods may further comprise administering a therapeutically effective amount of an agent selected from the group consisting of: insulin, insulin-sensitizing agents, insulin secretogogues, and combinations thereof. Preferably, the insulin-sensitizing agent is selected from the group consisting of biguanides, metformin, thiazolidinediones (glitazones), and combinations thereof. Preferably, the insulin secretogogue is selected from the group consisting of sulfonylureas, meglitinides, and combinations thereof (see, *e.g.*, Zangeneh et al., *Mayo Clin Proc.*, 78:471-479, 2003, incorporated by reference herein).

The present invention will now be illustrated by reference to the following examples which set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and are not to be construed as restricting the claimed invention in any way.

EXAMPLE I

(Materials and Methods)

Cell lines, transfections, expression vectors, western blots and antibodies

Cell lines. IRA-3T3 (3T3 cells transfected with a human insulin receptor cDNA have been previously described (Faria et al., *J. Biol. Chem.* 269:13922-13928 (1994)), and Herstatin-expressing MCF-7 cell clones were obtained using previously described methods (Shamieh et al., *FEBS Letters*, 568:163-166, 2004).

Transfections. For transient transfections, 2 μg of empty vector or 2 μg expression vector are added with Lipofectamine™ (GIBCO-BRL) to cells in 6 cm plates.

Western blot analysis, and antibodies. For Western blot analyses, whole-cell lysates or immunoprecipitated proteins were resolved by SDS-PAGE and transferred onto nitrocellulose membranes (BioRad, Hercules, CA). Blots were blocked in 5% milk and incubated with primary antibody overnight at 4°C. The antibodies included anti-insulin receptor (IR; against the β subunit), anti-IGF-IR, anti-IRS-1, anti-IRS-2, anti-phosphotyrosine, anti-phospho-Akt, anti-Akt, anti-phospho-ERK, anti-ERK, and anti-Shc antibodies (Santa Cruz Biotechnology, Transduction Laboratories, Cell Signaling Technologies, Upstate Laboratories, or Biosource). After washing, the blots were incubated with secondary antibody conjugated to HRP for 30 min (BioRad, Hercules, CA). The membranes were developed with SuperSignal™ West Dura (Pierce, Rockford, IL) and exposed to x-ray film.

Expression and purification of intron 8-encoded peptide (Int8) and Herstatin:

Receptor binding domain (RBD). Intron 8 cDNA, in the pET 30 bacterial expression vector (Novagen , Madison, WI), is expressed in bacteria (BL-21), and purified by nickel affinity chromatography as described (Doherty et al., *supra*).

Herstatin. For purification of insect Herstatin, S2 insect cells, stably transfected with 6xHis tagged-Herstatin in the pMT/BiP expression plasmid (Invitrogen, Carlsbad, CA), were induced with 100 μM cupric sulfate for about 16hrs. Herstatin was purified to about 90% purity by Ni-NTA (Qiagen, Valencia, CA) affinity chromatography as previously described (Jhabvala-Romero et al. *Supra.*).

Cell binding studies:

ELISA. Monolayer cultures of ~2 x 10^6 cells were plated in 6-well tissue culture plates, and were incubated with purified Herstatin for 2 hours at 4°C in serum-free DMEM. Cells were washed with Phosphate Buffered Saline (PBS) and extracted in 50mM Tris·HCl, pH 7.0, 1.0% NP-40. Herstatin bound to cells were quantified using a sandwich Herstatin ELISA per manufacturer's instructions (Upstate Biotechnology, Lake Placid, NY).

The dissociation constant (K_D) and maximal binding (B_{max}) of Herstatin were determined by nonlinear regression analysis of the plot of pmol of bound *versus* nM of Herstatin added. Statistical comparisons between different binding curves were performed by extra sums-of-squares F-test nonlinear regression coefficients. All tests were performed ($\alpha = 0.05$) using GraphPad™ Prism 4™ software (GraphPad™ Software, 1994-2003).

Pull-down assays with int8 peptide immobilized on protein S agarose:

About 100 μl of a 50% suspension of S-protein agarose (Novagen) is incubated with or without 100 μg of int8 peptide with an S-protein tag, at room temperature for 1hr, and then washed twice with 500 μl PBS. The agarose samples are then incubated at room temperature for 1 hr with 200 μg of transfected cell extract, then washed twice with 500 μl of PBS with 1% NP40. The proteins associated with the resin are eluted at 92°C for 2 min in 40μl of SDS-sample buffer, and analyzed as a Western blot.

Growth assays:

Cells ($4x10^4$) were plated in quadruplicate in 24-well plates, incubated in serum-free DMEM for 24 hours, and treated with either 10 nM insulin (Sigma) or an equivalent volume of vehicle (25 mM HEPES). At the indicated time points, cell monolayers were washed with PBS and incubated for 30 minutes at 37°C with 30 μl of MTS reagent [3-(4,5-dimethylthiazol-2-yl)-

5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl-2H-tetrazolium) inner salt Aqueous One Solution (Promega; Madison, WI) dissolved in 270 ml PBS] per well. Absorbance at 490 nm was determined a Bio-Tek plate reader.

EGFR inhibitor studies

Control MCF-7 cells were serum-starved overnight and treated with the EGFR kinase inhibitor AG1478 (Sigma) or vehicle (DMSO) for 5 minutes prior to the addition of 14 nM EGF or 10 nM insulin (Sigma). After growth factor treatment, cell lysates were prepared and analyzed for ERK and Akt/PKB activation as described above. The 24-hour treatment was done in regular growth medium.

EXAMPLE II

(Herstatin was shown to bind specifically to insulin receptor (IR) with nM binding affinity)

The interaction of Herstatin with IR in transfected 3T3 cells (IRA-3T3) was investigated. Herstatin bound specifically to IR at nM concentrations, and IR was thus shown herein to be a target of Herstatin.

Methods. Cell lines, expression vectors, protein purification, pull down assays, antibodies, Western blot analysis and ELISA assays were as described under EXAMPLE I, herein above.

Results. The interaction between Herstatin and IR was investigated. FIGURE 1 shows that Herstatin, purified from transfected S2 insect cells, exhibited dose-dependent binding to IR at nM concentrations. Increasing concentrations of Herstatin, expressed and purified from stably-transfected S2 insect cells, were added to 3T3 parental cells (filled triangles; "NIH-3T3")or 3T3 cells transfected with a human IR cDNA (filled squares; "IRA-3T3") as previously described (Shamieh et al., *FEBS Letters*, 568:163-166, 2004). After incubation for 2hrs on ice, the cells were washed twice with PBS, and the bound Herstatin was quantified using a Herstatin ELISA (Upstate). The data are plotted as Herstatin ELISA units versus concentration added. The results indicate that Herstatin binds at nM concentrations to cells expressing IR, but not to 3T3 parental cells.

These results demonstrate that Herstatin binds specifically to IR with nM binding affinity and that IGF-IR is a target of Herstatin.

EXAMPLE III

(Herstatin up-regulated insulin receptor (IR) expression, and activation

of IR by insulin in MCF-7 cells)

According to particular embodiments of the present invention, Herstatin not only up-regulates IR expression, but also up-regulates activation of IR by insulin (FIGURE 2).

Methods. Cell lines, expression vectors, protein purification, pull down assays, antibodies, Western blot analysis and ELISA assays were as described under EXAMPLE I, herein above. Insulin was added either to MCF-7 breast carcinoma cells, or to an MCF-7 cell line stably transfected with a Herstatin expression vector, to determine whether Herstatin expression affects IR expression, and/or insulin-stimulated IR signal transduction.

Results. FIGURE 2 shows that Herstatin expression not only up-regulated IR expression (including pro-IR), but also up-regulated IR activation (and thus signaling) in MCF-7 cells. Control and Herstatin-expressing MCF-7 cells were grown in complete medium prior to an overnight incubation in serum-free medium. Insulin was then added to the control and Herstatin-expressing cells and whole-cell lysates were prepared at the indicated times and processed directly for Western immunoblots with anti-insulin receptor (IR), phospho-Akt, Akt, phospho-ERK, and ERK antibodies, or first immunoprecipitated with anti-IR antibody and immunoprecipitates (IP) then analyzed by Western immunoblotting with anti-phosphotyrosine and anti-IR antibodies after transfer to nitrocellulose membranes. Following incubation of blots with primary antibodies, immunoreactive proteins were detected by enhanced chemiluminescence after a secondary incubation with HRP-conjugated secondary antisera. Similar results were obtained with a second Herstatin-expressing MCF-7 clone.

These results demonstrate that Herstatin not only up-regulates IR expression (including pro-IR), but also modulates IR-mediated signaling.

Additionally, as shown in FIGURE 2 (see also FIGURE 3 below), Herstatin up-regulated insulin-stimulated ERK activation (increased phospho-ERK).

EXAMPLE IV

(Herstatin expression amplified insulin-stimulated ERK activation in MCF-7 cells)

The effect of Herstatin expression on insulin-stimulated ERK activation/signaling was further investigated.

Methods. Methods were as described above under EXAMPLE III herein above.

Results. FIGURE 3 shows, in MCF-7 cells, that Herstatin expression amplified insulin-stimulated ERK activation. Control and Herstatin-expressing MCF-7 cells were treated and analyzed as those of Figure 2. Film exposures of enhanced chemiluminescence signals were

quantified by scanning densitometry, and the values for the phospho-ERK signals were normalized to the ERK signals to determine the relative level of ERK phosphorylation as a measure of activation.

Herstatin expression substantially amplified insulin-stimulated ERK activation in MCF-7 cells.

According to particular aspects of the present invention, this result supports a substantial utility for Herstatin in treating insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof.

This is because the MEK (MAPK kinase)-ERK pathway has been shown to be significantly involved in glucose transport (*e.g.*, Harmon et al., *Am. J. Physiol. Endocrinol. Metab.*, 287:E758-E766, 2004). Specifically, Harmon et al show specific inhibition of MAPK kinase (MEK) by the inhibitors PD-98059 and U-0216, resulting in significant inhibition of insulin-stimulated glucose uptake. The data support the importance of MEK for activation of GLUT4, and further, since the only target of MEK is ERK, the importance of the MEK (MAPK kinase)-ERK pathway for glucose transport.

EXAMPLE V

(Herstatin altered the expression of an array of proteins that are directly involved in insulin action.)

In addition to the regulation of insulin receptor protein, the regulation of the IRS-1 and IRS-2 proteins and Shc (that function as adapter proteins linking the activated insulin receptor to some of its downstream pathways), the expression of ERK and Akt/PKB, and the regulation of the IGF-IR (which may contribute to enhanced insulin receptor activation by decreasing the proportion of insulin receptor/IGF-I receptor hybrids, which do not respond to insulin) was investigated.

Methods. Cell lines, expression vectors, protein purification, antibodies and ELISA assays were as described under EXAMPLE I, herein above.

Results. Figure 4 shows that Herstatin altered the expression of an array of proteins that are directly involved in insulin action. Lysates from control and Herstatin-expressing MCF-7 cells were prepared from respective untreated (no insulin) cells following overnight incubation in serum-free media, and processed directly or (in the case of the IR) also immunoprecipitated prior to Western immunoblot analysis as described in relation to Figure 2.

These data illustrate that Herstatin: up-regulates insulin receptor protein as assessed by direct Western immunoblot and following immunoprecipitation; mediates the apparent phosphorylation state of the IRS-1 and IRS-2 (differentially down-regulated compared with IRS-1) proteins that function as adapter proteins linking the activated insulin receptor to some of its downstream pathways (see, *e.g.*, Le Roith 7 Zick, *Diabetes Care* 24:588-597, 2001, discussing role of IRS (IR substrate) proteins in IR-mediated signal transduction); elicits a slight decrease in IRS-2 expression; alters the relative expression of Shc isoforms expressed; increases the relative expression ratio of ERK1 and ERK2; and down-regulates the IGF-IR, which may contribute to enhanced insulin receptor activation by decreasing the proportion of IR/IGF-IR hybrids, which do not respond to insulin.

EXAMPLE VI

(The EGFR inhibitor AS1478 does not affect insulin signaling or lead to an increase in IR)

Figure 5 shows, according to particular aspects, that the EGFR inhibitor AS1478 did not affect insulin signaling.

Figure 6 shows, according to particular aspects, that inhibition of the EGF receptor with an EGF receptor-specific inhibitor did not lead to an increase in insulin receptor.

Other references of interest:

Jhabvala-Romero, F., A. Evans, S. Guo, M. Denton, and G. M. Clinton, 2003, Herstatin inhibits heregulin-mediated breast cancer cell growth and overcomes tamoxifen resistance in breast cancer cells that overexpress HER-2: *Oncogene*, v. 22, p. 8178-86.

Lin, Y. Z., S. W. Li, and G. M. Clinton, 1990, Insulin and epidermal growth factor stimulate phosphorylation of p185HER-2 in the breast carcinoma cell line, BT474: *Mol Cell Endocrinol*, v. 69, p. 111-9.

Christianson, T. A., J. K. Doherty, Y. J. Lin, E. E. Ramsey, R. Holmes, E. J. Keenan, and G. M. Clinton, 1998, NH2-terminally truncated HER-2/neu protein: relationship with shedding of the extracellular domain and with prognostic factors in breast cancer: *Cancer Res*, v. 58, p. 5123-9.

Garrett, T. P., N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H. J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoyne, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward, 2002, Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha: *Cell*, v. 110, p. 763-73.

Filmus, J., M. N. Pollak, J. G. Cairncross, and R. N. Buick, 1985, Amplified, overexpressed and rearranged epidermal growth factor receptor gene in a human astrocytoma cell line: *Biochem Biophys Res Commun*, v. 131, p. 207-15.

Filmus, J., M. N. Pollak, R. Cailleau, and R. N. Buick, 1985, MDA-468, a human breast cancer cell line with a high number of epidermal growth factor (EGF) *receptors, has an amplified EGF receptor gene and is growth inhibited by EGF: Biochem Biophys Res Commun*, v. 128, p. 898-905.

CLAIMS

1. A method for treating a condition associated with altered insulin receptor expression or altered insulin receptor-mediated signaling, said method comprising administering to a subject in need thereof, a therapeutically effective amount of Herstatin, or a variant thereof, that binds to the insulin receptor.

2. A method for treating a condition associated with altered insulin receptor expression or altered insulin receptor-mediated signaling, comprising administering to a subject in need thereof, a therapeutically effective amount of a Int8 RBD polypeptide, or a variant thereof, that binds to the insulin receptor.

3. The method of any one of claims 1 or 2, wherein the condition is at least one selected from the group consisting of insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, and neurodegenerative disorders.

4. The method of any one of claims 1 or 2, wherein the cell further expresses at least one target receptor selected from the group consisting of: EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4); and IGF-IR.

5. The method of claim 1, wherein the Herstatin, or variant thereof, comprises a polypeptide selected from the group consisting of SEQ ID NO:2, or a fragment of SEQ ID NO:2 of about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids are present, wherein at least one N-linked glycosylation site is present, and wherein the polypeptide binds to the insulin receptor.

6. The method of claim 1, wherein the Herstatin, or variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:32-42.

7. The method of claim 1, wherein the Herstatin, or variant thereof, comprises SEQ ID NO:32.

8. The method of claim 2, wherein the Int8 RBD polypeptide, or a variant thereof comprises a polypeptide selected from the group consisting of SEQ ID NO:1, or a fragment of SEQ ID NO:1 of about 50 to 79 contiguous residues in length, wherein the polypeptide binds to the insulin receptor.

9. The method of claim 2, wherein the Int8 RBD polypeptide, or a variant thereof, comprises a sequence selected from the group consisting of SEQ ID NOS:21-31,

10. The method of claim 2, wherein the Int8 RBD polypeptide, or a variant thereof, comprises SEQ ID NO:21.

11. The method of any one of claims 1 or 2, further comprising administering a therapeutically effective amount of a receptor-specific antibody that binds to a target receptor selected from the group consisting of: insulin receptor (IR), EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4), and IGF-IR.

12. The method of any one of claims 1 or 2, further comprising administration of a therapeutically effective amount of an agent selected from the group consisting of: insulin, insulin-sensitizing agents, insulin secretogogues, and combinations thereof.

13. The method of claim 12, wherein the insulin-sensitizing agent is selected from the group consisting of biguanides, metformin, thiazolidinediones (glitazones), and combinations thereof.

14. The method of claim 12, wherein the insulin secretogogue is selected from the group consisting of sulfonylureas, meglitinides, and combinations thereof.

15. A pharmaceutical composition for treating a condition associated with altered insulin receptor expression or altered insulin receptor-mediated signaling, comprising, Herstatin, or a variant thereof, that binds to the insulin receptor and a pharmaceutically acceptable carrier or excipient.

16. A pharmaceutical composition for treating a condition associated with altered insulin receptor expression or altered insulin receptor-mediated signaling, comprising, a Int8

RBD polypeptide, or a variant thereof, that binds to the insulin receptor and a pharmaceutically acceptable carrier or excipient.

17. The pharmaceutical composition of any one of claims 15 or 16, wherein the condition is selected from the group consisting of insulin resistance syndrome, pre-diabetic conditions, metabolic syndrome, type 1 and type 2 diabetes, cardiac disease, diabetes-associated vascular disease, atherosclerosis, hypertension, diabetes-associated lipid metabolism disorders (dyslipidemia), obesity, critical illness, neurodegenerative disorders, and combinations thereof.

18. The pharmaceutical composition of claim 15, wherein the Herstatin, or variant thereof, comprises a polypeptide selected from the group consisting of SEQ ID NO:2, or a fragment of SEQ ID NO:2 of about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids are present, wherein at least one N-linked glycosylation site is present, and wherein the polypeptide binds to the insulin receptor.

19. The pharmaceutical composition of claim 16, wherein the Int8 RBD polypeptide, or a variant thereof comprises a polypeptide selected from the group consisting of SEQ ID NO:1, or a fragment of SEQ ID NO:1 of about 50 to 79 contiguous residues in length, wherein the polypeptide binds to the insulin receptor.

20. The pharmaceutical composition of any one of claims 15 or 16, further comprising an agent selected from the group consisting of: insulin, insulin-sensitizing agents, insulin secretogogues, and combinations thereof.

21. The pharmaceutical composition of claim 20, wherein the insulin-sensitizing agent is selected from the group consisting of biguanides, metformin, thiazolidinediones (glitazones), and combinations thereof.

22. The pharmaceutical composition of claim 20, wherein the insulin secretogogue is selected from the group consisting of sulfonylureas, meglitinides, and combinations thereof..

23. A method for targeting a therapeutic agent to a cell expressing insulin receptor, comprising attaching the therapeutic agent to Herstatin, or to a variant thereof, that binds to the extracellular domain of a cellular target insulin receptor.

24. A method for targeting a therapeutic agent to a cell expressing insulin receptor, comprising attaching the therapeutic agent to a Int8 RBD polypeptide, or a variant thereof, that binds to the cellular target insulin receptor.

25. The method of any one of claims 23 or 24, wherein the cell further expresses a target receptor selected from the group consisting of: EGFR (HER-1, erbB-1); ΔEGFR; HER-2 (erbB-2); HER-3 (erbB-3); HER-4 (erbB-4); IGF-IR, and combinations thereof.

26. The method of claim 23, wherein the wherein the Herstatin, or variant thereof, comprises a polypeptide selected from the group consisting of SEQ ID NO:2, or a fragment of SEQ ID NO:2 of about 80 to 419 contiguous residues in length, wherein the C-terminal 79 contiguous amino acids are present, wherein at least one N-linked glycosylation site is present, and wherein the polypeptide binds to the insulin receptor.

27. The method of claim 24, wherein the Int8 RBD polypeptide, or a variant thereof comprises a polypeptide selected from the group consisting of SEQ ID NO:1, or a fragment of SEQ ID NO:1 of about 50 to 79 contiguous residues in length, wherein the polypeptide binds to the insulin receptor.

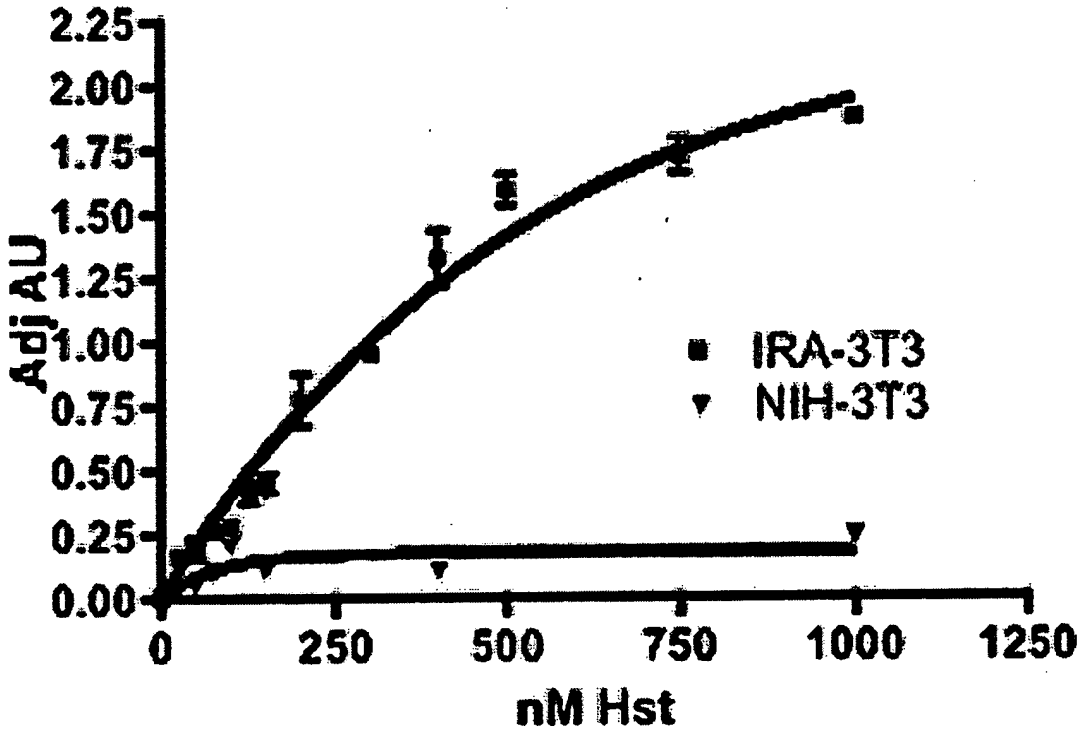

2.25
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Adj AU
0
250
500
750
1000
1250
nM Hst
IRA-3T3
NIH-3T3

FIG. 1

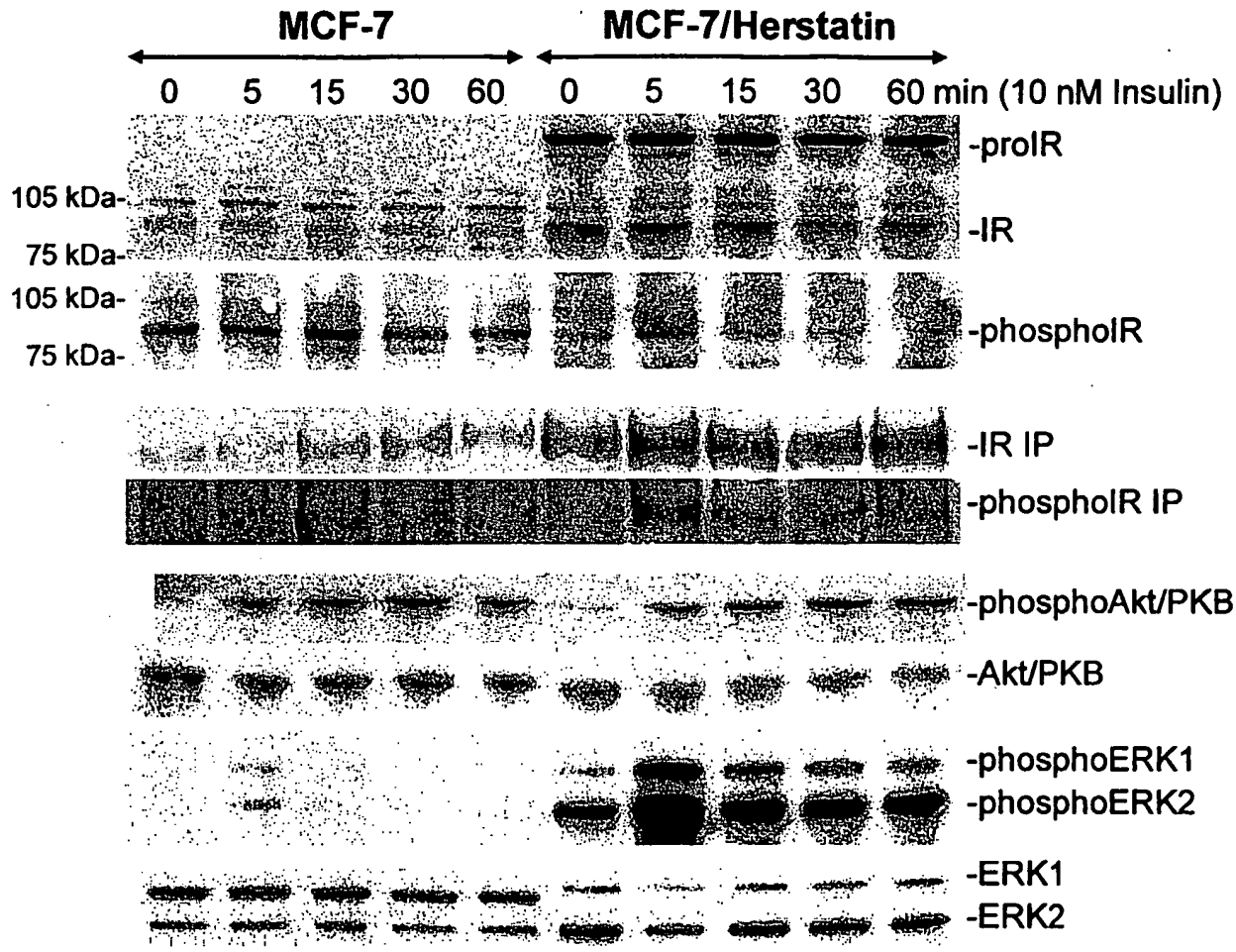

MCF-7
MCF-7/Herstatin
0 5 15 30 60 0 5 15 30 60 min (10 nM Insulin)
105 kDa-
75 kDa-
105 kDa-
75 kDa-
-proIR
-IR
-phosphoIR
-IR IP
-phosphoIR IP
-phosphoAkt/PKB
-Akt/PKB
-phosphoERK1
-phosphoERK2
-ERK1
-ERK2

FIG 2A

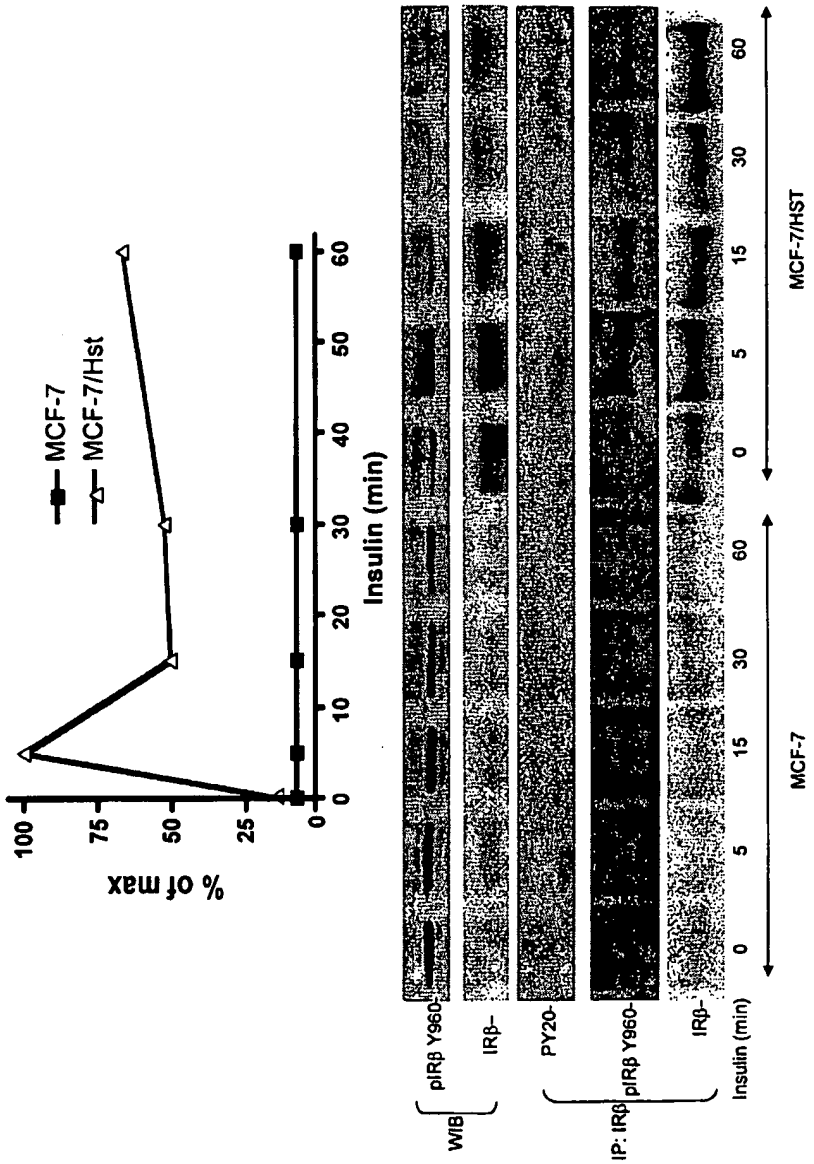

MCF-7
MCF-7/Hst
% of max
100
75
50
25
0
0
10
20
30
40
50
60
Insulin (min)
WIB
pIRβ Y960-
IRβ-
IP: IRβ
PY20-
pIRβ Y960-
IRβ-
Insulin (min)
0
5
15
30
60
MCF-7
0
5
15
30
60
MCF-7/HST

FIG. 2B

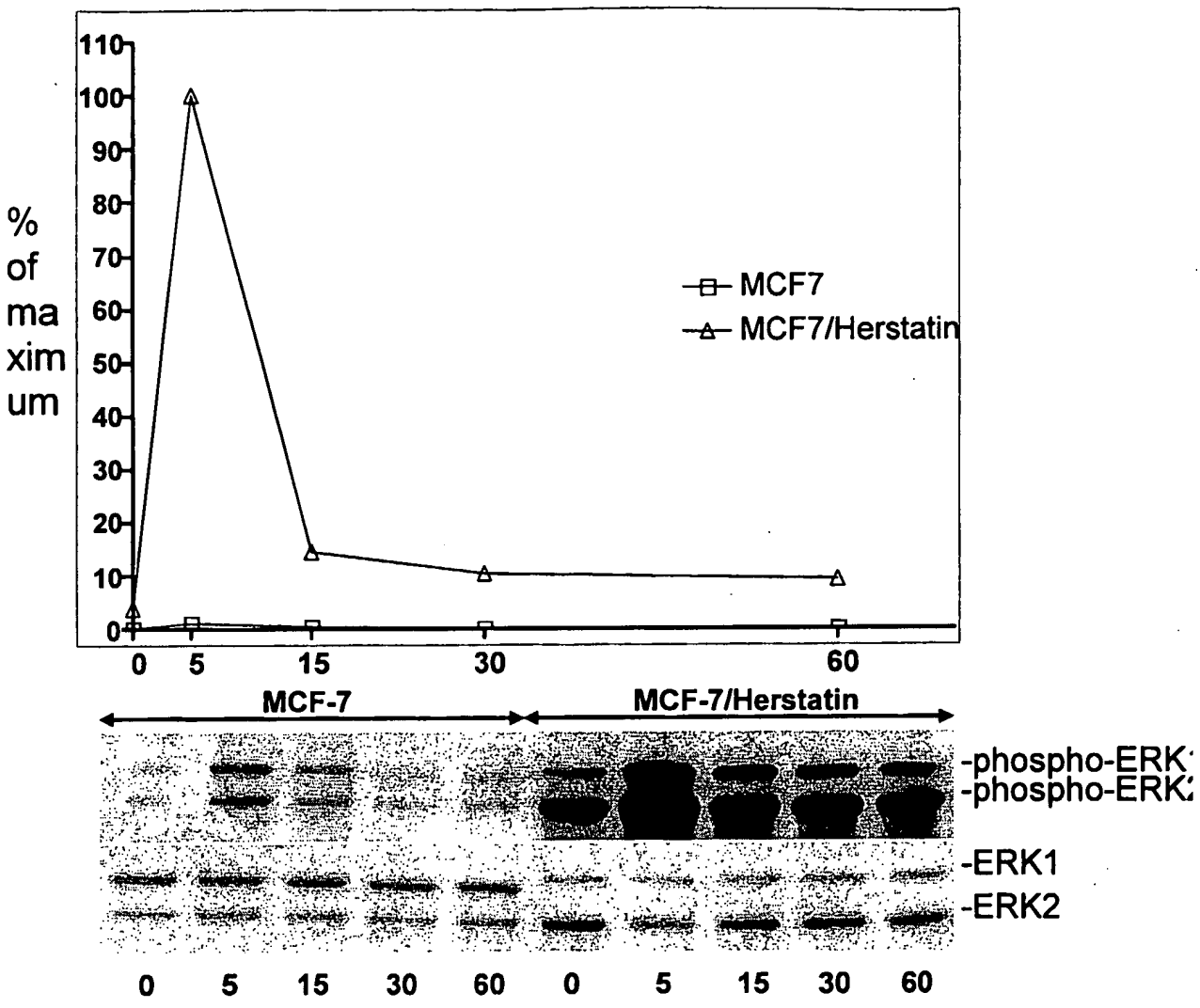

% of maximum
110
100
90
80
70
60
50
40
30
20
10
0
0 5 15 30 60
MCF7
MCF7/Herstatin
MCF-7
MCF-7/Herstatin
-phospho-ERK1
-phospho-ERK2
-ERK1
-ERK2
0 5 15 30 60 0 5 15 30 60

FIG. 3A

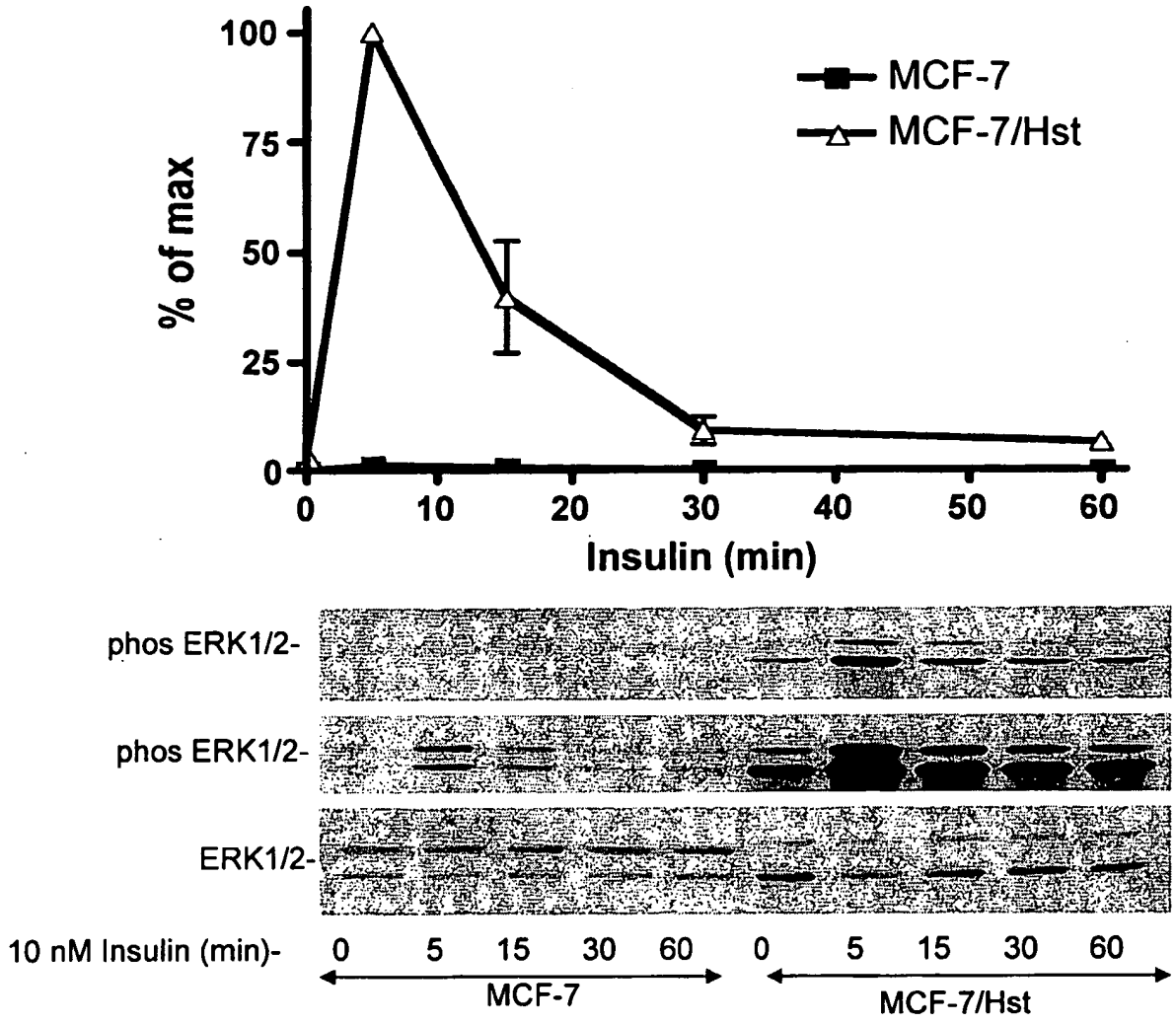

100
75
50
25
0
% of max
MCF-7
MCF-7/Hst
0
10
20
30
40
50
60
Insulin (min)
phos ERK1/2-
phos ERK1/2-
ERK1/2-
10 nM Insulin (min)-
0 5 15 30 60
0 5 15 30 60
MCF-7
MCF-7/Hst

FIG. 3B

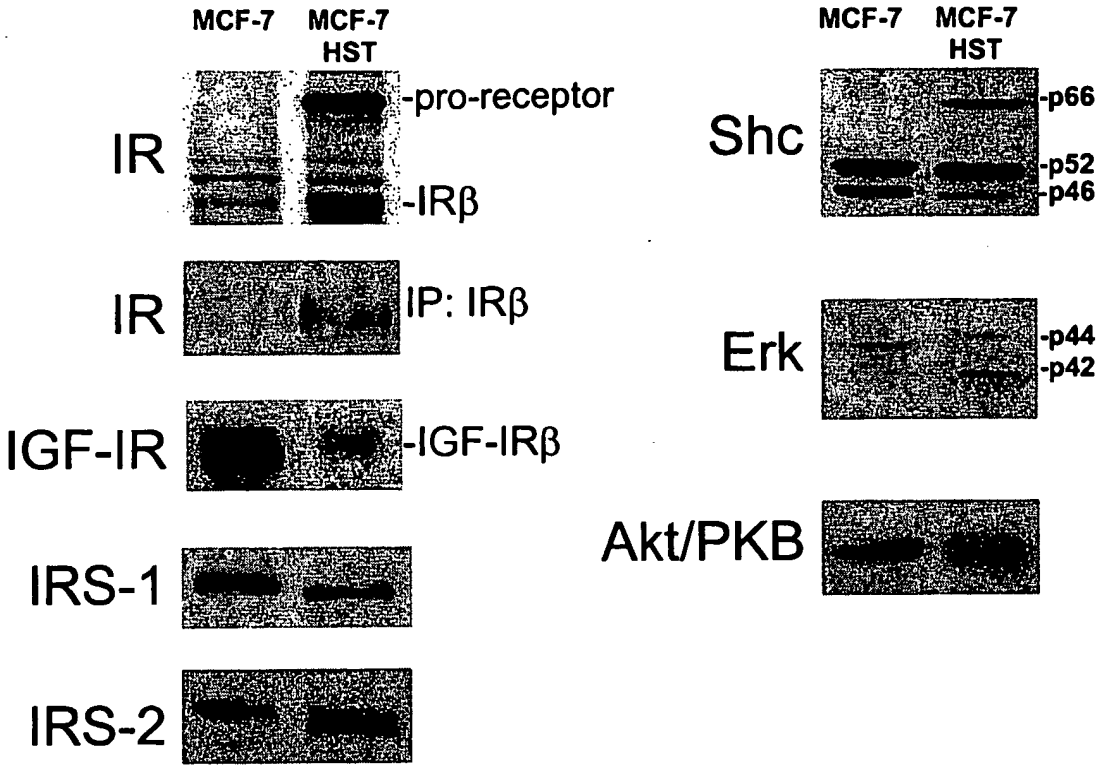

MCF-7
MCF-7 HST
IR
-pro-receptor
-IRβ
IR
IP: IRβ
IGF-IR
-IGF-IRβ
IRS-1
IRS-2
MCF-7
MCF-7 HST
Shc
-p66
-p52
-p46
Erk
-p44
-p42
Akt/PKB

FIG. 4A

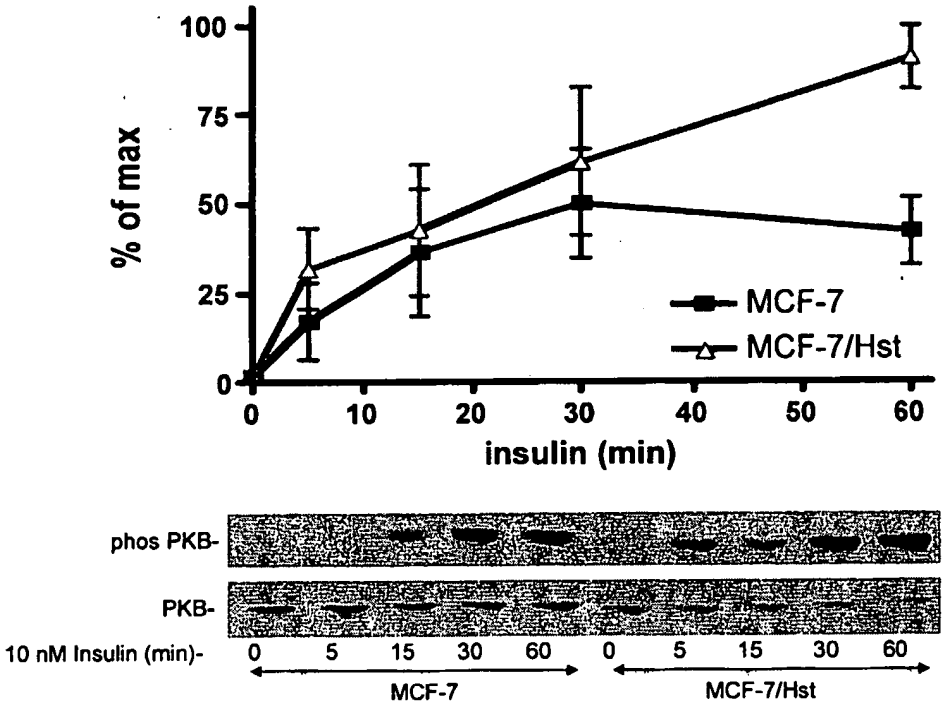
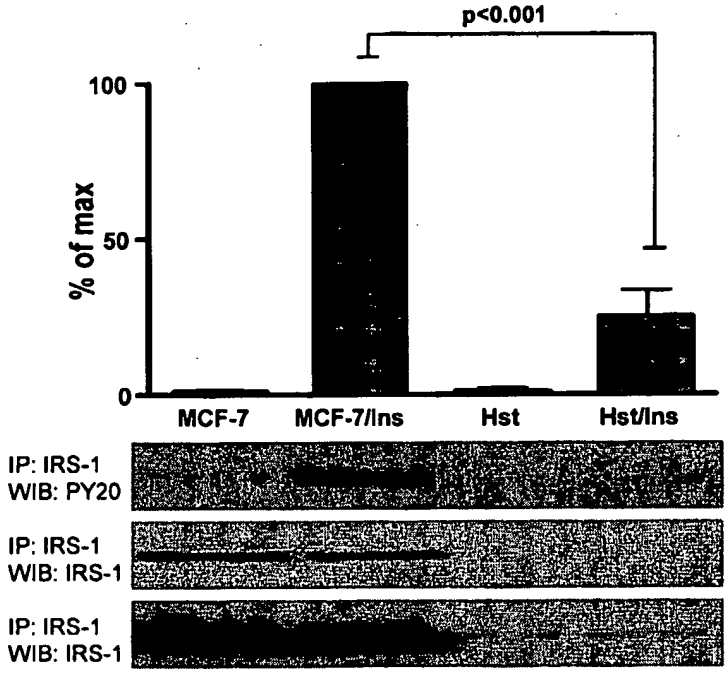
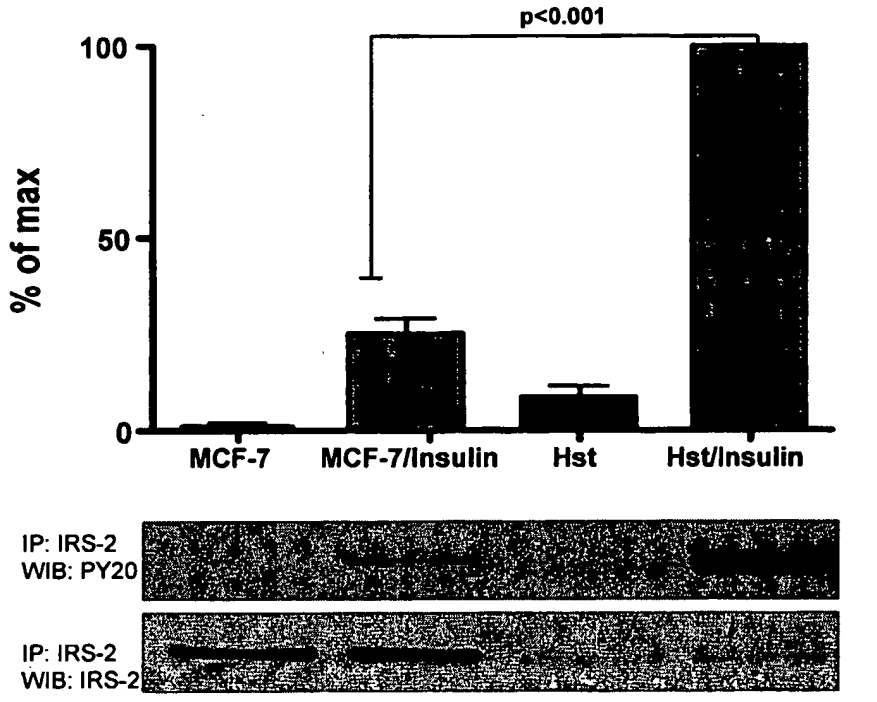

100
75
50
25
0
% of max
0
10
20
30
40
50
60
insulin (min)
MCF-7
MCF-7/Hst
phos PKB-
PKB-
10 nM Insulin (min)-
0
5
15
30
60
0
5
15
30
60
MCF-7
MCF-7/Hst

FIG. 4B



* Not corrected for differences in total IRS-1 between cell lines

FIG. 4C



* Not corrected for differences in total IRS-2 between cell lines

FIG. 4D

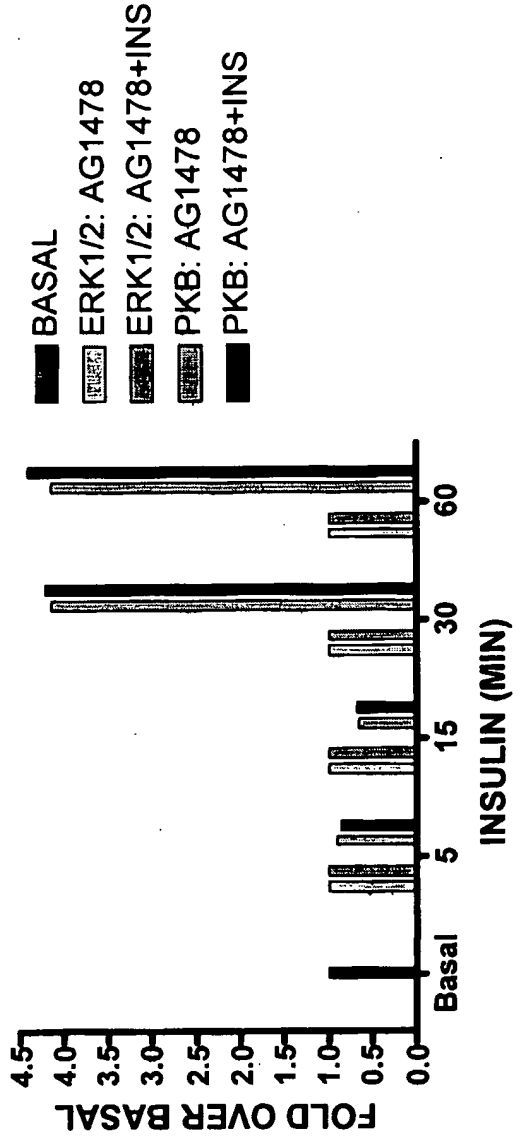

BASAL
ERK1/2: AG1478
ERK1/2: AG1478+INS
PKB: AG1478
PKB: AG1478+INS
FOLD OVER BASAL
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Basal
5
15
30
60
INSULIN (MIN)

FIG. 5

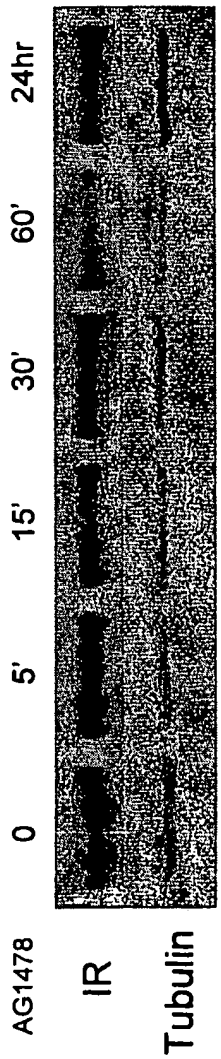


FIG. 6

SEQUENCE LISTING

<110> Oregon Health & Science University
Clinton, Gail
Roberts, Charles T

<120> COMPOSITIONS AND METHODS FOR TREATING DISEASE

<130> 49321-146

<150> US 60/616,596
<151> 2004-10-05

<150> US 60/688,355
<151> 2005-06-06

<160> 42

<170> PatentIn version 3.3

<210> 1
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa reflects Thr or Ser variants

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa reflects Leu or Pro variants

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa reflects Pro or Leu variants

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> Xaa reflects Leu or Gln variants

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> Xaa reflects Arg or Cys variants

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> Xaa reflects Met or Leu variants

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> Xaa reflects Gly, Asp, Ala or Val variants

<220>
<221> MISC_FEATURE
<222> (31)..(31)

<223> Xaa reflects Arg or Ile variants

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> Xaa reflects Leu or Ile variants

<220>
<221> MISC_FEATURE
<222> (54)..(54)
<223> Xaa reflects Pro or Arg variants

<220>
<221> MISC_FEATURE
<222> (64)..(64)
<223> Xaa reflects Pro or Leu variants

<220>
<221> MISC_FEATURE
<222> (73)..(73)
<223> Xaa reflects Asp or Asn variants

<400> 1

Gly Xaa His Ser Xaa Xaa Pro Arg Pro Ala Ala Val Pro Val Pro Xaa
1 5 10 15

Xaa Xaa Gln Pro Xaa Pro Ala His Pro Val Leu Ser Phe Leu Xaa Pro
20 25 30

Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val Gly Arg Gly Xaa
50 55 60

Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 2
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (342)..(342)
<223> Xaa reflects Thr or Ser variants

<220>
<221> MISC_FEATURE
<222> (345)..(345)
<223> Xaa reflects Leu or Pro variants

<220>
<221> MISC_FEATURE
<222> (346)..(346)
<223> Xaa reflects Pro or Leu variants

<220>
<221> MISC_FEATURE
<222> (356)..(356)
<223> Xaa reflects Leu or Gln variants

<220>
<221> MISC_FEATURE
<222> (357)..(357)
<223> Xaa reflects Arg or Cys variants

<220>
<221> MISC_FEATURE
<222> (358)..(358)
<223> Xaa reflects Met or Leu variants

<220>
<221> MISC_FEATURE
<222> (361)..(361)
<223> Xaa reflects Gly, Asp, Ala or Val variants

<220>
<221> MISC_FEATURE
<222> (371)..(371)
<223> Xaa reflects Arg or Ile variants

<220>
<221> MISC_FEATURE
<222> (376)..(376)
<223> Xaa reflects Leu or Ile variants

<220>
<221> MISC_FEATURE
<222> (394)..(394)
<223> Xaa reflects Pro or Arg variants

<220>
<221> MISC_FEATURE
<222> (404)..(404)
<223> Xaa reflects Pro or Leu variants

<220>
<221> MISC_FEATURE
<222> (413)..(413)
<223> Xaa reflects Asp or Asn variants

<400> 2

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                  10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Cys Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60
```

```
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Cys Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Arg Arg Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Cys Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Cys Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Cys Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Lys Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320
```

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Xaa His Ser Xaa Xaa Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Xaa Xaa Xaa Gln Pro Xaa Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Xaa Pro Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Asp Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Xaa Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 3
<211> 79
<212> PRT
<213> artificial sequence

<220>
<223> ECDIIIA domain non-binding mutant (Arg to Ile mutation at residue 31)

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (5)..(6)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (16)..(16)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (18)..(18)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature
<222> (36)..(36)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (54)..(54)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (64)..(64)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (73)..(73)
<223> Xaa can be any naturally occurring amino acid

<400> 3

Gly Xaa His Ser Xaa Xaa Pro Arg Pro Ala Ala Val Pro Val Pro Xaa
1 5 10 15

Arg Xaa Gln Pro Xaa Pro Ala His Pro Val Leu Ser Phe Leu Ile Pro
20 25 30

Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val Gly Arg Gly Xaa
50 55 60

Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 4
<211> 419
<212> PRT
<213> artificial sequence

<220>
<223> Herstatin; receptor non-binding mutant (Arg to Ile mutation at residue 371)

<220>
<221> misc_feature
<222> (342)..(342)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (345)..(346)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (356)..(356)

<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (358)..(358)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (361)..(361)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (376)..(376)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (394)..(394)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (404)..(404)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (413)..(413)
<223> Xaa can be any naturally occurring amino acid

<400> 4

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                  10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Cys Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Cys Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125
```

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Cys Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Cys Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Cys Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Lys Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Xaa His Ser Xaa Xaa Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Xaa Arg Xaa Gln Pro Xaa Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Ile Pro Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

```
Leu Ala Pro Leu Asp Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val
385                 390                 395                 400

Gly Arg Gly Xaa Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg
                405                 410                 415

Tyr Glu Gly
```

<210> 5
<211> 5616
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (247)..(3879)
<223> HER-1 coding sequence

<400> 5

```
ccccggcgca gcgcggccgc agcagcctcc gccccccgca cggtgtgagc gcccgacgcg     60

gccgaggcgg ccggagtccc gagctagccc cggcggccgc cgccgcccag accggacgac    120

aggccacctc gtcggcgtcc gcccgagtcc ccgcctcgcc gccaacgcca caaccaccgc    180

gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga    240

gcagcg atg cga ccc tcc ggg acg gcc ggg gca gcg ctc ctg gcg ctg        288
       Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu
       1               5                   10

ctg gct gcg ctc tgc ccg gcg agt cgg gct ctg gag gaa aag aaa gtt      336
Leu Ala Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val
15                  20                  25                  30

tgc caa ggc acg agt aac aag ctc acg cag ttg ggc act ttt gaa gat      384
Cys Gln Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp
                35                  40                  45

cat ttt ctc agc ctc cag agg atg ttc aat aac tgt gag gtg gtc ctt      432
His Phe Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu
            50                  55                  60

ggg aat ttg gaa att acc tat gtg cag agg aat tat gat ctt tcc ttc      480
Gly Asn Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe
        65                  70                  75

tta aag acc atc cag gag gtg gct ggt tat gtc ctc att gcc ctc aac      528
Leu Lys Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn
    80                  85                  90

aca gtg gag cga att cct ttg gaa aac ctg cag atc atc aga gga aat      576
Thr Val Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn
95                  100                 105                 110

atg tac tac gaa aat tcc tat gcc tta gca gtc tta tct aac tat gat      624
Met Tyr Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp
```

```
                115                 120                 125

gca aat aaa acc gga ctg aag gag ctg ccc atg aga aat tta cag gaa      672
Ala Asn Lys Thr Gly Leu Lys Glu Leu Pro Met Arg Asn Leu Gln Glu
            130                 135                 140

atc ctg cat ggc gcc gtg cgg ttc agc aac aac cct gcc ctg tgc aac      720
Ile Leu His Gly Ala Val Arg Phe Ser Asn Asn Pro Ala Leu Cys Asn
        145                 150                 155

gtg gag agc atc cag tgg cgg gac ata gtc agc agt gac ttt ctc agc      768
Val Glu Ser Ile Gln Trp Arg Asp Ile Val Ser Ser Asp Phe Leu Ser
    160                 165                 170

aac atg tcg atg gac ttc cag aac cac ctg ggc agc tgc caa aag tgt      816
Asn Met Ser Met Asp Phe Gln Asn His Leu Gly Ser Cys Gln Lys Cys
175                 180                 185                 190

gat cca agc tgt ccc aat ggg agc tgc tgg ggt gca gga gag gag aac      864
Asp Pro Ser Cys Pro Asn Gly Ser Cys Trp Gly Ala Gly Glu Glu Asn
                195                 200                 205

tgc cag aaa ctg acc aaa atc atc tgt gcc cag cag tgc tcc ggg cgc      912
Cys Gln Lys Leu Thr Lys Ile Ile Cys Ala Gln Gln Cys Ser Gly Arg
            210                 215                 220

tgc cgt ggc aag tcc ccc agt gac tgc tgc cac aac cag tgt gct gca      960
Cys Arg Gly Lys Ser Pro Ser Asp Cys Cys His Asn Gln Cys Ala Ala
        225                 230                 235

ggc tgc aca ggc ccc cgg gag agc gac tgc ctg gtc tgc cgc aaa ttc     1008
Gly Cys Thr Gly Pro Arg Glu Ser Asp Cys Leu Val Cys Arg Lys Phe
    240                 245                 250

cga gac gaa gcc acg tgc aag gac acc tgc ccc cca ctc atg ctc tac     1056
Arg Asp Glu Ala Thr Cys Lys Asp Thr Cys Pro Pro Leu Met Leu Tyr
255                 260                 265                 270

aac ccc acc acg tac cag atg gat gtg aac ccc gag ggc aaa tac agc     1104
Asn Pro Thr Thr Tyr Gln Met Asp Val Asn Pro Glu Gly Lys Tyr Ser
                275                 280                 285

ttt ggt gcc acc tgc gtg aag aag tgt ccc cgt aat tat gtg gtg aca     1152
Phe Gly Ala Thr Cys Val Lys Lys Cys Pro Arg Asn Tyr Val Val Thr
            290                 295                 300

gat cac ggc tcg tgc gtc cga gcc tgt ggg gcc gac agc tat gag atg     1200
Asp His Gly Ser Cys Val Arg Ala Cys Gly Ala Asp Ser Tyr Glu Met
        305                 310                 315

gag gaa gac ggc gtc cgc aag tgt aag aag tgc gaa ggg cct tgc cgc     1248
Glu Glu Asp Gly Val Arg Lys Cys Lys Lys Cys Glu Gly Pro Cys Arg
    320                 325                 330

aaa gtg tgt aac gga ata ggt att ggt gaa ttt aaa gac tca ctc tcc     1296
Lys Val Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp Ser Leu Ser
335                 340                 345                 350

ata aat gct acg aat att aaa cac ttc aaa aac tgc acc tcc atc agt     1344
Ile Asn Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr Ser Ile Ser
                355                 360                 365

ggc gat ctc cac atc ctg ccg gtg gca ttt agg ggt gac tcc ttc aca     1392
Gly Asp Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp Ser Phe Thr
```

```
            370                 375                 380

cat act cct cct ctg gat cca cag gaa ctg gat att ctg aaa acc gta     1440
His Thr Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu Lys Thr Val
        385                 390                 395

aag gaa atc aca ggg ttt ttg ctg att cag gct tgg cct gaa aac agg     1488
Lys Glu Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro Glu Asn Arg
    400                 405                 410

acg gac ctc cat gcc ttt gag aac cta gaa atc ata cgc ggc agg acc     1536
Thr Asp Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg Gly Arg Thr
415                 420                 425                 430

aag caa cat ggt cag ttt tct ctt gca gtc gtc agc ctg aac ata aca     1584
Lys Gln His Gly Gln Phe Ser Leu Ala Val Val Ser Leu Asn Ile Thr
                435                 440                 445

tcc ttg gga tta cgc tcc ctc aag gag ata agt gat gga gat gtg ata     1632
Ser Leu Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly Asp Val Ile
            450                 455                 460

att tca gga aac aaa aat ttg tgc tat gca aat aca ata aac tgg aaa     1680
Ile Ser Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile Asn Trp Lys
        465                 470                 475

aaa ctg ttt ggg acc tcc ggt cag aaa acc aaa att ata agc aac aga     1728
Lys Leu Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile Ser Asn Arg
    480                 485                 490

ggt gaa aac agc tgc aag gcc aca ggc cag gtc tgc cat gcc ttg tgc     1776
Gly Glu Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His Ala Leu Cys
495                 500                 505                 510

tcc ccc gag ggc tgc tgg ggc ccg gag ccc agg gac tgc gtc tct tgc     1824
Ser Pro Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys Val Ser Cys
                515                 520                 525

cgg aat gtc agc cga ggc agg gaa tgc gtg gac aag tgc aac ctt ctg     1872
Arg Asn Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys Asn Leu Leu
            530                 535                 540

gag ggt gag cca agg gag ttt gtg gag aac tct gag tgc ata cag tgc     1920
Glu Gly Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys Ile Gln Cys
        545                 550                 555

cac cca gag tgc ctg cct cag gcc atg aac atc acc tgc aca gga cgg     1968
His Pro Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys Thr Gly Arg
    560                 565                 570

gga cca gac aac tgt atc cag tgt gcc cac tac att gac ggc ccc cac     2016
Gly Pro Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp Gly Pro His
575                 580                 585                 590

tgc gtc aag acc tgc ccg gca gga gtc atg gga gaa aac aac acc ctg     2064
Cys Val Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn Asn Thr Leu
                595                 600                 605

gtc tgg aag tac gca gac gcc ggc cat gtg tgc cac ctg tgc cat cca     2112
Val Trp Lys Tyr Ala Asp Ala Gly His Val Cys His Leu Cys His Pro
            610                 615                 620

aac tgc acc tac gga tgc act ggg cca ggt ctt gaa ggc tgt cca acg     2160
Asn Cys Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly Cys Pro Thr
```

```
        625                 630                 635

aat ggg cct aag atc ccg tcc atc gcc act ggg atg gtg ggg gcc ctc     2208
Asn Gly Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala Leu
    640                 645                 650

ctc ttg ctg ctg gtg gtg gcc ctg ggg atc ggc ctc ttc atg cga agg     2256
Leu Leu Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met Arg Arg
655                 660                 665                 670

cgc cac atc gtt cgg aag cgc acg ctg cgg agg ctg ctg cag gag agg     2304
Arg His Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu Gln Glu Arg
                675                 680                 685

gag ctt gtg gag cct ctt aca ccc agt gga gaa gct ccc aac caa gct     2352
Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro Asn Gln Ala
            690                 695                 700

ctc ttg agg atc ttg aag gaa act gaa ttc aaa aag atc aaa gtg ctg     2400
Leu Leu Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile Lys Val Leu
        705                 710                 715

ggc tcc ggt gcg ttc ggc acg gtg tat aag gga ctc tgg atc cca gaa     2448
Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp Ile Pro Glu
    720                 725                 730

ggt gag aaa gtt aaa att ccc gtc gct atc aag gaa tta aga gaa gca     2496
Gly Glu Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala
735                 740                 745                 750

aca tct ccg aaa gcc aac aag gaa atc ctc gat gaa gcc tac gtg atg     2544
Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met
                755                 760                 765

gcc agc gtg gac aac ccc cac gtg tgc cgc ctg ctg ggc atc tgc ctc     2592
Ala Ser Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu
            770                 775                 780

acc tcc acc gtg cag ctc atc acg cag ctc atg ccc ttc ggc tgc ctc     2640
Thr Ser Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu
        785                 790                 795

ctg gac tat gtc cgg gaa cac aaa gac aat att ggc tcc cag tac ctg     2688
Leu Asp Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu
    800                 805                 810

ctc aac tgg tgt gtg cag atc gca aag ggc atg aac tac ttg gag gac     2736
Leu Asn Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp
815                 820                 825                 830

cgt cgc ttg gtg cac cgc gac ctg gca gcc agg aac gta ctg gtg aaa     2784
Arg Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys
                835                 840                 845

aca ccg cag cat gtc aag atc aca gat ttt ggg ctg gcc aaa ctg ctg     2832
Thr Pro Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu
            850                 855                 860

ggt gcg gaa gag aaa gaa tac cat gca gaa gga ggc aaa gtg cct atc     2880
Gly Ala Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile
        865                 870                 875

aag tgg atg gca ttg gaa tca att tta cac aga atc tat acc cac cag     2928
Lys Trp Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln
```

880 885 890

agt gat gtc tgg agc tac ggg gtg acc gtt tgg gag ttg atg acc ttt 2976
Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe
895 900 905 910

gga tcc aag cca tat gac gga atc cct gcc agc gag atc tcc tcc atc 3024
Gly Ser Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile
915 920 925

ctg gag aaa gga gaa cgc ctc cct cag cca ccc ata tgt acc atc gat 3072
Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp
930 935 940

gtc tac atg atc atg gtc aag tgc tgg atg ata gac gca gat agt cgc 3120
Val Tyr Met Ile Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg
945 950 955

cca aag ttc cgt gag ttg atc atc gaa ttc tcc aaa atg gcc cga gac 3168
Pro Lys Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met Ala Arg Asp
960 965 970

ccc cag cgc tac ctt gtc att cag ggg gat gaa aga atg cat ttg cca 3216
Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met His Leu Pro
975 980 985 990

agt cct aca gac tcc aac ttc tac cgt gcc ctg atg gat gaa gaa gac 3264
Ser Pro Thr Asp Ser Asn Phe Tyr Arg Ala Leu Met Asp Glu Glu Asp
995 1000 1005

atg gac gac gtg gtg gat gcc gac gag tac ctc atc cca cag cag 3309
Met Asp Asp Val Val Asp Ala Asp Glu Tyr Leu Ile Pro Gln Gln
1010 1015 1020

ggc ttc ttc agc agc ccc tcc acg tca cgg act ccc ctc ctg agc 3354
Gly Phe Phe Ser Ser Pro Ser Thr Ser Arg Thr Pro Leu Leu Ser
1025 1030 1035

tct ctg agt gca acc agc aac aat tcc acc gtg gct tgc att gat 3399
Ser Leu Ser Ala Thr Ser Asn Asn Ser Thr Val Ala Cys Ile Asp
1040 1045 1050

aga aat ggg ctg caa agc tgt ccc atc aag gaa gac agc ttc ttg 3444
Arg Asn Gly Leu Gln Ser Cys Pro Ile Lys Glu Asp Ser Phe Leu
1055 1060 1065

cag cga tac agc tca gac ccc aca ggc gcc ttg act gag gac agc 3489
Gln Arg Tyr Ser Ser Asp Pro Thr Gly Ala Leu Thr Glu Asp Ser
1070 1075 1080

ata gac gac acc ttc ctc cca gtg cct gaa tac ata aac cag tcc 3534
Ile Asp Asp Thr Phe Leu Pro Val Pro Glu Tyr Ile Asn Gln Ser
1085 1090 1095

gtt ccc aaa agg ccc gct ggc tct gtg cag aat cct gtc tat cac 3579
Val Pro Lys Arg Pro Ala Gly Ser Val Gln Asn Pro Val Tyr His
1100 1105 1110

aat cag cct ctg aac ccc gcg ccc agc aga gac cca cac tac cag 3624
Asn Gln Pro Leu Asn Pro Ala Pro Ser Arg Asp Pro His Tyr Gln
1115 1120 1125

gac ccc cac agc act gca gtg ggc aac ccc gag tat ctc aac act 3669
Asp Pro His Ser Thr Ala Val Gly Asn Pro Glu Tyr Leu Asn Thr

```
            1130                1135                1140

gtc cag ccc acc  tgt gtc aac agc aca  ttc gac agc cct gcc  cac      3714
Val Gln Pro Thr  Cys Val Asn Ser Thr  Phe Asp Ser Pro Ala  His
            1145                1150                1155

tgg gcc cag aaa  ggc agc cac caa att  agc ctg gac aac cct  gac      3759
Trp Ala Gln Lys  Gly Ser His Gln Ile  Ser Leu Asp Asn Pro  Asp
            1160                1165                1170

tac cag cag gac  ttc ttt ccc aag gaa  gcc aag cca aat ggc  atc      3804
Tyr Gln Gln Asp  Phe Phe Pro Lys Glu  Ala Lys Pro Asn Gly  Ile
            1175                1180                1185

ttt aag ggc tcc  aca gct gaa aat gca  gaa tac cta agg gtc  gcg      3849
Phe Lys Gly Ser  Thr Ala Glu Asn Ala  Glu Tyr Leu Arg Val  Ala
            1190                1195                1200

cca caa agc agt  gaa ttt att gga gca  tga ccacggagga tagtatgagc     3899
Pro Gln Ser Ser  Glu Phe Ile Gly Ala
            1205                1210

cctaaaaatc cagactcttt cgatacccag gaccaagcca cagcaggtcc tccatcccaa   3959

cagccatgcc cgcattagct cttagaccca cagactggtt ttgcaacgtt tacaccgact   4019

agccaggaag tacttccacc tcgggcacat tttgggaagt tgcattcctt tgtcttcaaa   4079

ctgtgaagca tttacagaaa cgcatccagc aagaatattg tccctttgag cagaaattta   4139

tctttcaaag aggtatattt gaaaaaaaaa aaaagtatat gtgaggattt ttattgattg   4199

gggatcttgg agtttttcat tgtcgctatt gatttttact tcaatgggct cttccaacaa   4259

ggaagaagct tgctggtagc acttgctacc ctgagttcat ccaggcccaa ctgtgagcaa   4319

ggagcacaag ccacaagtct tccagaggat gcttgattcc agtggttctg cttcaaggct   4379

tccactgcaa aacactaaag atccaagaag gccttcatgg ccccagcagg ccggatcggt   4439

actgtatcaa gtcatggcag gtacagtagg ataagccact ctgtcccttc ctgggcaaag   4499

aagaaacgga ggggatggaa ttcttcctta gacttacttt tgtaaaaatg tccccacggt   4559

acttactccc cactgatgga ccagtggttt ccagtcatga gcgttagact gacttgtttg   4619

tcttccattc cattgttttg aaactcagta tgctgcccct gtcttgctgt catgaaatca   4679

gcaagagagg atgacacatc aaataataac tcggattcca gcccacattg gattcatcag   4739

catttggacc aatagcccac agctgagaat gtggaatacc taaggatagc accgcttttg   4799

ttctcgcaaa aacgtatctc ctaatttgag gctcagatga aatgcatcag gtcctttggg   4859

gcatagatca gaagactaca aaaatgaagc tgctctgaaa tctcctttag ccatcacccc   4919

aacccccaa aattagtttg tgttacttat ggaagatagt tttctccttt tacttcactt   4979

caaaagcttt ttactcaaag agtatatgtt ccctccaggt cagctgcccc caaacccccc   5039

ccttacgctt tgtcacacaa aaagtgtctc tgccttgagt catctattca agcacttaca   5099

gctctggcca caacagggca ttttacaggt gcgaatgaca gtagcattat gagtagtgtg   5159
```

gaattcaggt agtaaatatg aaactagggt ttgaaattga taatgctttc acaacatttg 5219

cagatgtttt agaaggaaaa aagttccttc ctaaaataat ttctctacaa ttggaagatt 5279

ggaagattca gctagttagg agcccacctt ttttcctaat ctgtgtgtgc cctgtaacct 5339

gactggttaa cagcagtcct ttgtaaacag tgttttaaac tctcctagtc aatatccacc 5399

ccatccaatt tatcaaggaa gaaatggttc agaaaatatt ttcagcctac agttatgttc 5459

agtcacacac acatacaaaa tgttcctttt gcttttaaag taatttttga ctcccagatc 5519

agtcagagcc cctacagcat tgttaagaaa gtatttgatt tttgtctcaa tgaaaataaa 5579

actatattca tttccactct aaaaaaaaaa aaaaaaa 5616

<210> 6
<211> 1210
<212> PRT
<213> Homo sapiens

<400> 6

Met Arg Pro Ser Gly Thr Ala Gly Ala Ala Leu Leu Ala Leu Leu Ala
1 5 10 15

Ala Leu Cys Pro Ala Ser Arg Ala Leu Glu Glu Lys Lys Val Cys Gln
20 25 30

Gly Thr Ser Asn Lys Leu Thr Gln Leu Gly Thr Phe Glu Asp His Phe
35 40 45

Leu Ser Leu Gln Arg Met Phe Asn Asn Cys Glu Val Val Leu Gly Asn
50 55 60

Leu Glu Ile Thr Tyr Val Gln Arg Asn Tyr Asp Leu Ser Phe Leu Lys
65 70 75 80

Thr Ile Gln Glu Val Ala Gly Tyr Val Leu Ile Ala Leu Asn Thr Val
85 90 95

Glu Arg Ile Pro Leu Glu Asn Leu Gln Ile Ile Arg Gly Asn Met Tyr
100 105 110

Tyr Glu Asn Ser Tyr Ala Leu Ala Val Leu Ser Asn Tyr Asp Ala Asn
115 120 125

Lys Thr Gly Leu Lys Glu Leu Pro Met Arg Asn Leu Gln Glu Ile Leu
130 135 140

His Gly Ala Val Arg Phe Ser Asn Asn Pro Ala Leu Cys Asn Val Glu
145 150 155 160

Ser Ile Gln Trp Arg Asp Ile Val Ser Ser Asp Phe Leu Ser Asn Met
165 170 175

Ser Met Asp Phe Gln Asn His Leu Gly Ser Cys Gln Lys Cys Asp Pro
180 185 190

Ser Cys Pro Asn Gly Ser Cys Trp Gly Ala Gly Glu Glu Asn Cys Gln
195 200 205

Lys Leu Thr Lys Ile Ile Cys Ala Gln Gln Cys Ser Gly Arg Cys Arg
210 215 220

Gly Lys Ser Pro Ser Asp Cys Cys His Asn Gln Cys Ala Ala Gly Cys
225 230 235 240

Thr Gly Pro Arg Glu Ser Asp Cys Leu Val Cys Arg Lys Phe Arg Asp
245 250 255

Glu Ala Thr Cys Lys Asp Thr Cys Pro Pro Leu Met Leu Tyr Asn Pro
260 265 270

Thr Thr Tyr Gln Met Asp Val Asn Pro Glu Gly Lys Tyr Ser Phe Gly
275 280 285

Ala Thr Cys Val Lys Lys Cys Pro Arg Asn Tyr Val Val Thr Asp His
290 295 300

Gly Ser Cys Val Arg Ala Cys Gly Ala Asp Ser Tyr Glu Met Glu Glu
305 310 315 320

Asp Gly Val Arg Lys Cys Lys Lys Cys Glu Gly Pro Cys Arg Lys Val
325 330 335

Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp Ser Leu Ser Ile Asn
340 345 350

Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr Ser Ile Ser Gly Asp
355 360 365

Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp Ser Phe Thr His Thr
370 375 380

Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu Lys Thr Val Lys Glu
385 390 395 400

Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro Glu Asn Arg Thr Asp
405 410 415

Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg Gly Arg Thr Lys Gln
420 425 430

His Gly Gln Phe Ser Leu Ala Val Val Ser Leu Asn Ile Thr Ser Leu
435 440 445

Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly Asp Val Ile Ile Ser
450 455 460

Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile Asn Trp Lys Lys Leu
465 470 475 480

Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile Ser Asn Arg Gly Glu
485 490 495

Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His Ala Leu Cys Ser Pro
500 505 510

Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys Val Ser Cys Arg Asn
515 520 525

Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys Asn Leu Leu Glu Gly
530 535 540

Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys Ile Gln Cys His Pro
545 550 555 560

Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys Thr Gly Arg Gly Pro
565 570 575

Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp Gly Pro His Cys Val
580 585 590

Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn Asn Thr Leu Val Trp
595 600 605

Lys Tyr Ala Asp Ala Gly His Val Cys His Leu Cys His Pro Asn Cys
610 615 620

Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly Cys Pro Thr Asn Gly
625 630 635 640

Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val Gly Ala Leu Leu Leu
645 650 655

Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe Met Arg Arg Arg His
660 665 670

Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu Gln Glu Arg Glu Leu
675 680 685

Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro Asn Gln Ala Leu Leu
690 695 700

Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile Lys Val Leu Gly Ser
705 710 715 720

Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp Ile Pro Glu Gly Glu
725 730 735

Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg Glu Ala Thr Ser
740 745 750

Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Ser
755 760 765

Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser
770 775 780

Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu Asp
785 790 795 800

Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn
805 810 815

Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg
820 825 830

Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Thr Pro
835 840 845

Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala Lys Leu Leu Gly Ala
850 855 860

Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys Val Pro Ile Lys Trp
865 870 875 880

Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr His Gln Ser Asp
885 890 895

Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ser
900 905 910

Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu
915 920 925

```
Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr
    930                 935                 940

Met Ile Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro Lys
945                 950                 955                 960

Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met Ala Arg Asp Pro Gln
                965                 970                 975

Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met His Leu Pro Ser Pro
            980                 985                 990

Thr Asp Ser Asn Phe Tyr Arg Ala  Leu Met Asp Glu Glu  Asp Met Asp
        995                 1000                1005

Asp Val  Val Asp Ala Asp Glu  Tyr Leu Ile Pro Gln  Gln Gly Phe
    1010                1015                1020

Phe Ser  Ser Pro Ser Thr Ser  Arg Thr Pro Leu Leu  Ser Ser Leu
    1025                1030                1035

Ser Ala  Thr Ser Asn Asn Ser  Thr Val Ala Cys Ile  Asp Arg Asn
    1040                1045                1050

Gly Leu  Gln Ser Cys Pro Ile  Lys Glu Asp Ser Phe  Leu Gln Arg
    1055                1060                1065

Tyr Ser  Ser Asp Pro Thr Gly  Ala Leu Thr Glu Asp  Ser Ile Asp
    1070                1075                1080

Asp Thr  Phe Leu Pro Val Pro  Glu Tyr Ile Asn Gln  Ser Val Pro
    1085                1090                1095

Lys Arg  Pro Ala Gly Ser Val  Gln Asn Pro Val Tyr  His Asn Gln
    1100                1105                1110

Pro Leu  Asn Pro Ala Pro Ser  Arg Asp Pro His Tyr  Gln Asp Pro
    1115                1120                1125

His Ser  Thr Ala Val Gly Asn  Pro Glu Tyr Leu Asn  Thr Val Gln
    1130                1135                1140

Pro Thr  Cys Val Asn Ser Thr  Phe Asp Ser Pro Ala  His Trp Ala
    1145                1150                1155

Gln Lys  Gly Ser His Gln Ile  Ser Leu Asp Asn Pro  Asp Tyr Gln
    1160                1165                1170
```

```
Gln Asp  Phe Phe Pro Lys Glu  Ala Lys Pro Asn Gly  Ile Phe Lys
    1175                 1180                 1185

Gly Ser  Thr Ala Glu Asn Ala  Glu Tyr Leu Arg Val  Ala Pro Gln
    1190                 1195                 1200

Ser Ser  Glu Phe Ile Gly Ala
    1205                 1210
```

<210> 7
<211> 4815
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (247)..(3078)
<223> delta EGFR coding sequence [represents in-frame deletion of 801 bp (275-1027) of the ECD of EGFR corresponding to exons 2-7 of the gene]

<400> 7

```
ccccggcgca gcgcggccgc agcagcctcc gccccccgca cggtgtgagc gcccgacgcg     60

gccgaggcgg ccggagtccc gagctagccc cggcggccgc cgccgcccag accggacgac    120

aggccacctc gtcggcgtcc gcccgagtcc ccgcctcgcc gccaacgcca caaccaccgc    180

gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga    240

gcagcg atg cga ccc tcc ggg acg gcc ggg gca gtg gat gtg aac ccc        288
       Met Arg Pro Ser Gly Thr Ala Gly Ala Val Asp Val Asn Pro
       1               5                  10

gag ggc aaa tac agc ttt ggt gcc acc tgc gtg aag aag tgt ccc cgt      336
Glu Gly Lys Tyr Ser Phe Gly Ala Thr Cys Val Lys Lys Cys Pro Arg
15                  20                  25                  30

aat tat gtg gtg aca gat cac ggc tcg tgc gtc cga gcc tgt ggg gcc      384
Asn Tyr Val Val Thr Asp His Gly Ser Cys Val Arg Ala Cys Gly Ala
                35                  40                  45

gac agc tat gag atg gag gaa gac ggc gtc cgc aag tgt aag aag tgc      432
Asp Ser Tyr Glu Met Glu Glu Asp Gly Val Arg Lys Cys Lys Lys Cys
            50                  55                  60

gaa ggg cct tgc cgc aaa gtg tgt aac gga ata ggt att ggt gaa ttt      480
Glu Gly Pro Cys Arg Lys Val Cys Asn Gly Ile Gly Ile Gly Glu Phe
        65                  70                  75

aaa gac tca ctc tcc ata aat gct acg aat att aaa cac ttc aaa aac      528
Lys Asp Ser Leu Ser Ile Asn Ala Thr Asn Ile Lys His Phe Lys Asn
    80                  85                  90

tgc acc tcc atc agt ggc gat ctc cac atc ctg ccg gtg gca ttt agg      576
Cys Thr Ser Ile Ser Gly Asp Leu His Ile Leu Pro Val Ala Phe Arg
95                  100                 105                 110

ggt gac tcc ttc aca cat act cct cct ctg gat cca cag gaa ctg gat      624
Gly Asp Ser Phe Thr His Thr Pro Pro Leu Asp Pro Gln Glu Leu Asp
```

```
                115                 120                 125

att ctg aaa acc gta aag gaa atc aca ggg ttt ttg ctg att cag gct      672
Ile Leu Lys Thr Val Lys Glu Ile Thr Gly Phe Leu Leu Ile Gln Ala
            130                 135                 140

tgg cct gaa aac agg acg gac ctc cat gcc ttt gag aac cta gaa atc      720
Trp Pro Glu Asn Arg Thr Asp Leu His Ala Phe Glu Asn Leu Glu Ile
        145                 150                 155

ata cgc ggc agg acc aag caa cat ggt cag ttt tct ctt gca gtc gtc      768
Ile Arg Gly Arg Thr Lys Gln His Gly Gln Phe Ser Leu Ala Val Val
    160                 165                 170

agc ctg aac ata aca tcc ttg gga tta cgc tcc ctc aag gag ata agt      816
Ser Leu Asn Ile Thr Ser Leu Gly Leu Arg Ser Leu Lys Glu Ile Ser
175                 180                 185                 190

gat gga gat gtg ata att tca gga aac aaa aat ttg tgc tat gca aat      864
Asp Gly Asp Val Ile Ile Ser Gly Asn Lys Asn Leu Cys Tyr Ala Asn
                195                 200                 205

aca ata aac tgg aaa aaa ctg ttt ggg acc tcc ggt cag aaa acc aaa      912
Thr Ile Asn Trp Lys Lys Leu Phe Gly Thr Ser Gly Gln Lys Thr Lys
            210                 215                 220

att ata agc aac aga ggt gaa aac agc tgc aag gcc aca ggc cag gtc      960
Ile Ile Ser Asn Arg Gly Glu Asn Ser Cys Lys Ala Thr Gly Gln Val
        225                 230                 235

tgc cat gcc ttg tgc tcc ccc gag ggc tgc tgg ggc ccg gag ccc agg     1008
Cys His Ala Leu Cys Ser Pro Glu Gly Cys Trp Gly Pro Glu Pro Arg
    240                 245                 250

gac tgc gtc tct tgc cgg aat gtc agc cga ggc agg gaa tgc gtg gac     1056
Asp Cys Val Ser Cys Arg Asn Val Ser Arg Gly Arg Glu Cys Val Asp
255                 260                 265                 270

aag tgc aac ctt ctg gag ggt gag cca agg gag ttt gtg gag aac tct     1104
Lys Cys Asn Leu Leu Glu Gly Glu Pro Arg Glu Phe Val Glu Asn Ser
                275                 280                 285

gag tgc ata cag tgc cac cca gag tgc ctg cct cag gcc atg aac atc     1152
Glu Cys Ile Gln Cys His Pro Glu Cys Leu Pro Gln Ala Met Asn Ile
            290                 295                 300

acc tgc aca gga cgg gga cca gac aac tgt atc cag tgt gcc cac tac     1200
Thr Cys Thr Gly Arg Gly Pro Asp Asn Cys Ile Gln Cys Ala His Tyr
        305                 310                 315

att gac ggc ccc cac tgc gtc aag acc tgc ccg gca gga gtc atg gga     1248
Ile Asp Gly Pro His Cys Val Lys Thr Cys Pro Ala Gly Val Met Gly
    320                 325                 330

gaa aac aac acc ctg gtc tgg aag tac gca gac gcc ggc cat gtg tgc     1296
Glu Asn Asn Thr Leu Val Trp Lys Tyr Ala Asp Ala Gly His Val Cys
335                 340                 345                 350

cac ctg tgc cat cca aac tgc acc tac gga tgc act ggg cca ggt ctt     1344
His Leu Cys His Pro Asn Cys Thr Tyr Gly Cys Thr Gly Pro Gly Leu
                355                 360                 365

gaa ggc tgt cca acg aat ggg cct aag atc ccg tcc atc gcc act ggg     1392
Glu Gly Cys Pro Thr Asn Gly Pro Lys Ile Pro Ser Ile Ala Thr Gly
```

```
          370                 375                 380

atg gtg ggg gcc ctc ctc ttg ctg ctg gtg gtg gcc ctg ggg atc ggc     1440
Met Val Gly Ala Leu Leu Leu Leu Leu Val Val Ala Leu Gly Ile Gly
        385                 390                 395

ctc ttc atg cga agg cgc cac atc gtt cgg aag cgc acg ctg cgg agg     1488
Leu Phe Met Arg Arg Arg His Ile Val Arg Lys Arg Thr Leu Arg Arg
    400                 405                 410

ctg ctg cag gag agg gag ctt gtg gag cct ctt aca ccc agt gga gaa     1536
Leu Leu Gln Glu Arg Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Glu
415                 420                 425                 430

gct ccc aac caa gct ctc ttg agg atc ttg aag gaa act gaa ttc aaa     1584
Ala Pro Asn Gln Ala Leu Leu Arg Ile Leu Lys Glu Thr Glu Phe Lys
                435                 440                 445

aag atc aaa gtg ctg ggc tcc ggt gcg ttc ggc acg gtg tat aag gga     1632
Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly
            450                 455                 460

ctc tgg atc cca gaa ggt gag aaa gtt aaa att ccc gtc gct atc aag     1680
Leu Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val Ala Ile Lys
        465                 470                 475

gaa tta aga gaa gca aca tct ccg aaa gcc aac aag gaa atc ctc gat     1728
Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp
    480                 485                 490

gaa gcc tac gtg atg gcc agc gtg gac aac ccc cac gtg tgc cgc ctg     1776
Glu Ala Tyr Val Met Ala Ser Val Asp Asn Pro His Val Cys Arg Leu
495                 500                 505                 510

ctg ggc atc tgc ctc acc tcc acc gtg cag ctc atc acg cag ctc atg     1824
Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr Gln Leu Met
                515                 520                 525

ccc ttc ggc tgc ctc ctg gac tat gtc cgg gaa cac aaa gac aat att     1872
Pro Phe Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys Asp Asn Ile
            530                 535                 540

ggc tcc cag tac ctg ctc aac tgg tgt gtg cag atc gca aag ggc atg     1920
Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala Lys Gly Met
        545                 550                 555

aac tac ttg gag gac cgt cgc ttg gtg cac cgc gac ctg gca gcc agg     1968
Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu Ala Ala Arg
    560                 565                 570

aac gta ctg gtg aaa aca ccg cag cat gtc aag atc aca gat ttt ggg     2016
Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe Gly
575                 580                 585                 590

ctg gcc aaa ctg ctg ggt gcg gaa gag aaa gaa tac cat gca gaa gga     2064
Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His Ala Glu Gly
                595                 600                 605

ggc aaa gtg cct atc aag tgg atg gca ttg gaa tca att tta cac aga     2112
Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu His Arg
            610                 615                 620

atc tat acc cac cag agt gat gtc tgg agc tac ggg gtg acc gtt tgg     2160
Ile Tyr Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp
```

625 630 635

gag ttg atg acc ttt gga tcc aag cca tat gac gga atc cct gcc agc 2208
Glu Leu Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile Pro Ala Ser
640 645 650

gag atc tcc tcc atc ctg gag aaa gga gaa cgc ctc cct cag cca ccc 2256
Glu Ile Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro
655 660 665 670

ata tgt acc atc gat gtc tac atg atc atg gtc aag tgc tgg atg ata 2304
Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile
675 680 685

gac gca gat agt cgc cca aag ttc cgt gag ttg atc atc gaa ttc tcc 2352
Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile Glu Phe Ser
690 695 700

aaa atg gcc cga gac ccc cag cgc tac ctt gtc att cag ggg gat gaa 2400
Lys Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp Glu
705 710 715

aga atg cat ttg cca agt cct aca gac tcc aac ttc tac cgt gcc ctg 2448
Arg Met His Leu Pro Ser Pro Thr Asp Ser Asn Phe Tyr Arg Ala Leu
720 725 730

atg gat gaa gaa gac atg gac gac gtg gtg gat gcc gac gag tac ctc 2496
Met Asp Glu Glu Asp Met Asp Asp Val Val Asp Ala Asp Glu Tyr Leu
735 740 745 750

atc cca cag cag ggc ttc ttc agc agc ccc tcc acg tca cgg act ccc 2544
Ile Pro Gln Gln Gly Phe Phe Ser Ser Pro Ser Thr Ser Arg Thr Pro
755 760 765

ctc ctg agc tct ctg agt gca acc agc aac aat tcc acc gtg gct tgc 2592
Leu Leu Ser Ser Leu Ser Ala Thr Ser Asn Asn Ser Thr Val Ala Cys
770 775 780

att gat aga aat ggg ctg caa agc tgt ccc atc aag gaa gac agc ttc 2640
Ile Asp Arg Asn Gly Leu Gln Ser Cys Pro Ile Lys Glu Asp Ser Phe
785 790 795

ttg cag cga tac agc tca gac ccc aca ggc gcc ttg act gag gac agc 2688
Leu Gln Arg Tyr Ser Ser Asp Pro Thr Gly Ala Leu Thr Glu Asp Ser
800 805 810

ata gac gac acc ttc ctc cca gtg cct gaa tac ata aac cag tcc gtt 2736
Ile Asp Asp Thr Phe Leu Pro Val Pro Glu Tyr Ile Asn Gln Ser Val
815 820 825 830

ccc aaa agg ccc gct ggc tct gtg cag aat cct gtc tat cac aat cag 2784
Pro Lys Arg Pro Ala Gly Ser Val Gln Asn Pro Val Tyr His Asn Gln
835 840 845

cct ctg aac ccc gcg ccc agc aga gac cca cac tac cag gac ccc cac 2832
Pro Leu Asn Pro Ala Pro Ser Arg Asp Pro His Tyr Gln Asp Pro His
850 855 860

agc act gca gtg ggc aac ccc gag tat ctc aac act gtc cag ccc acc 2880
Ser Thr Ala Val Gly Asn Pro Glu Tyr Leu Asn Thr Val Gln Pro Thr
865 870 875

tgt gtc aac agc aca ttc gac agc cct gcc cac tgg gcc cag aaa ggc 2928
Cys Val Asn Ser Thr Phe Asp Ser Pro Ala His Trp Ala Gln Lys Gly

880 885 890

agc cac caa att agc ctg gac aac cct gac tac cag cag gac ttc ttt 2976
Ser His Gln Ile Ser Leu Asp Asn Pro Asp Tyr Gln Gln Asp Phe Phe
895 900 905 910

ccc aag gaa gcc aag cca aat ggc atc ttt aag ggc tcc aca gct gaa 3024
Pro Lys Glu Ala Lys Pro Asn Gly Ile Phe Lys Gly Ser Thr Ala Glu
915 920 925

aat gca gaa tac cta agg gtc gcg cca caa agc agt gaa ttt att gga 3072
Asn Ala Glu Tyr Leu Arg Val Ala Pro Gln Ser Ser Glu Phe Ile Gly
930 935 940

gca tga ccacggagga tagtatgagc cctaaaaatc cagactcttt cgatacccag 3128
Ala

gaccaagcca cagcaggtcc tccatcccaa cagccatgcc cgcattagct cttagaccca 3188

cagactggtt ttgcaacgtt tacaccgact agccaggaag tacttccacc tcgggcacat 3248

tttgggaagt tgcattcctt tgtcttcaaa ctgtgaagca tttacagaaa cgcatccagc 3308

aagaatattg tccctttgag cagaaattta tctttcaaag aggtatattt gaaaaaaaaa 3368

aaaagtatat gtgaggattt ttattgattg gggatcttgg agtttttcat tgtcgctatt 3428

gatttttact tcaatgggct cttccaacaa ggaagaagct tgctggtagc acttgctacc 3488

ctgagttcat ccaggcccaa ctgtgagcaa ggagcacaag ccacaagtct tccagaggat 3548

gcttgattcc agtggttctg cttcaaggct tccactgcaa aacactaaag atccaagaag 3608

gccttcatgg ccccagcagg ccggatcggt actgtatcaa gtcatggcag gtacagtagg 3668

ataagccact ctgtcccttc ctgggcaaag aagaaacgga ggggatggaa ttcttcctta 3728

gacttacttt tgtaaaaatg tccccacggt acttactccc cactgatgga ccagtggttt 3788

ccagtcatga gcgttagact gacttgtttg tcttccattc cattgttttg aaactcagta 3848

tgctgcccct gtcttgctgt catgaaatca gcaagagagg atgacacatc aaataataac 3908

tcggattcca gcccacattg gattcatcag catttggacc aatagcccac agctgagaat 3968

gtggaatacc taaggatagc accgcttttg ttctcgcaaa aacgtatctc ctaatttgag 4028

gctcagatga aatgcatcag gtcctttggg gcatagatca gaagactaca aaaatgaagc 4088

tgctctgaaa tctcctttag ccatcacccc aacccccaa aattagtttg tgttacttat 4148

ggaagatagt tttctccttt tacttcactt caaaagcttt ttactcaaag agtatatgtt 4208

ccctccaggt cagctgcccc caaacccct ccttacgctt tgtcacacaa aaagtgtctc 4268

tgccttgagt catctattca agcacttaca gctctggcca caacagggca ttttacaggt 4328

gcgaatgaca gtagcattat gagtagtgtg gaattcaggt agtaaatatg aaactagggt 4388

ttgaaattga taatgctttc acaacatttg cagatgtttt agaaggaaaa aagttccttc 4448

ctaaaataat ttctctacaa ttggaagatt ggaagattca gctagttagg agcccacctt 4508

ttttcctaat ctgtgtgtgc cctgtaacct gactggttaa cagcagtcct ttgtaaacag 4568

tgttttaaac tctcctagtc aatatccacc ccatccaatt tatcaaggaa gaaatggttc 4628

agaaaatatt ttcagcctac agttatgttc agtcacacac acatacaaaa tgttcctttt 4688

gcttttaaag taatttttga ctcccagatc agtcagagcc cctacagcat tgttaagaaa 4748

gtatttgatt tttgtctcaa tgaaaataaa actatattca tttccactct aaaaaaaaaa 4808

aaaaaaa 4815

<210> 8
<211> 943
<212> PRT
<213> Homo sapiens

<400> 8

Met Arg Pro Ser Gly Thr Ala Gly Ala Val Asp Val Asn Pro Glu Gly
1 5 10 15

Lys Tyr Ser Phe Gly Ala Thr Cys Val Lys Lys Cys Pro Arg Asn Tyr
20 25 30

Val Val Thr Asp His Gly Ser Cys Val Arg Ala Cys Gly Ala Asp Ser
35 40 45

Tyr Glu Met Glu Glu Asp Gly Val Arg Lys Cys Lys Lys Cys Glu Gly
50 55 60

Pro Cys Arg Lys Val Cys Asn Gly Ile Gly Ile Gly Glu Phe Lys Asp
65 70 75 80

Ser Leu Ser Ile Asn Ala Thr Asn Ile Lys His Phe Lys Asn Cys Thr
85 90 95

Ser Ile Ser Gly Asp Leu His Ile Leu Pro Val Ala Phe Arg Gly Asp
100 105 110

Ser Phe Thr His Thr Pro Pro Leu Asp Pro Gln Glu Leu Asp Ile Leu
115 120 125

Lys Thr Val Lys Glu Ile Thr Gly Phe Leu Leu Ile Gln Ala Trp Pro
130 135 140

Glu Asn Arg Thr Asp Leu His Ala Phe Glu Asn Leu Glu Ile Ile Arg
145 150 155 160

Gly Arg Thr Lys Gln His Gly Gln Phe Ser Leu Ala Val Val Ser Leu
165 170 175

```
Asn Ile Thr Ser Leu Gly Leu Arg Ser Leu Lys Glu Ile Ser Asp Gly
            180                 185                 190

Asp Val Ile Ile Ser Gly Asn Lys Asn Leu Cys Tyr Ala Asn Thr Ile
        195                 200                 205

Asn Trp Lys Lys Leu Phe Gly Thr Ser Gly Gln Lys Thr Lys Ile Ile
    210                 215                 220

Ser Asn Arg Gly Glu Asn Ser Cys Lys Ala Thr Gly Gln Val Cys His
225                 230                 235                 240

Ala Leu Cys Ser Pro Glu Gly Cys Trp Gly Pro Glu Pro Arg Asp Cys
                245                 250                 255

Val Ser Cys Arg Asn Val Ser Arg Gly Arg Glu Cys Val Asp Lys Cys
            260                 265                 270

Asn Leu Leu Glu Gly Glu Pro Arg Glu Phe Val Glu Asn Ser Glu Cys
        275                 280                 285

Ile Gln Cys His Pro Glu Cys Leu Pro Gln Ala Met Asn Ile Thr Cys
    290                 295                 300

Thr Gly Arg Gly Pro Asp Asn Cys Ile Gln Cys Ala His Tyr Ile Asp
305                 310                 315                 320

Gly Pro His Cys Val Lys Thr Cys Pro Ala Gly Val Met Gly Glu Asn
                325                 330                 335

Asn Thr Leu Val Trp Lys Tyr Ala Asp Ala Gly His Val Cys His Leu
            340                 345                 350

Cys His Pro Asn Cys Thr Tyr Gly Cys Thr Gly Pro Gly Leu Glu Gly
        355                 360                 365

Cys Pro Thr Asn Gly Pro Lys Ile Pro Ser Ile Ala Thr Gly Met Val
    370                 375                 380

Gly Ala Leu Leu Leu Leu Leu Val Val Ala Leu Gly Ile Gly Leu Phe
385                 390                 395                 400

Met Arg Arg Arg His Ile Val Arg Lys Arg Thr Leu Arg Arg Leu Leu
                405                 410                 415

Gln Glu Arg Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Glu Ala Pro
            420                 425                 430
```

```
Asn Gln Ala Leu Leu Arg Ile Leu Lys Glu Thr Glu Phe Lys Lys Ile
        435                 440                 445

Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu Trp
    450                 455                 460

Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu
465                 470                 475                 480

Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala
                485                 490                 495

Tyr Val Met Ala Ser Val Asp Asn Pro His Val Cys Arg Leu Leu Gly
            500                 505                 510

Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe
        515                 520                 525

Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys Asp Asn Ile Gly Ser
    530                 535                 540

Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala Lys Gly Met Asn Tyr
545                 550                 555                 560

Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn Val
                565                 570                 575

Leu Val Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe Gly Leu Ala
            580                 585                 590

Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His Ala Glu Gly Gly Lys
        595                 600                 605

Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr
    610                 615                 620

Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu
625                 630                 635                 640

Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile
                645                 650                 655

Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys
            660                 665                 670

Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met Ile Asp Ala
        675                 680                 685
```

```
Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile Glu Phe Ser Lys Met
    690                 695                 700

Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp Glu Arg Met
705                 710                 715                 720

His Leu Pro Ser Pro Thr Asp Ser Asn Phe Tyr Arg Ala Leu Met Asp
                725                 730                 735

Glu Glu Asp Met Asp Asp Val Val Asp Ala Asp Glu Tyr Leu Ile Pro
            740                 745                 750

Gln Gln Gly Phe Phe Ser Ser Pro Ser Thr Ser Arg Thr Pro Leu Leu
        755                 760                 765

Ser Ser Leu Ser Ala Thr Ser Asn Asn Ser Thr Val Ala Cys Ile Asp
    770                 775                 780

Arg Asn Gly Leu Gln Ser Cys Pro Ile Lys Glu Asp Ser Phe Leu Gln
785                 790                 795                 800

Arg Tyr Ser Ser Asp Pro Thr Gly Ala Leu Thr Glu Asp Ser Ile Asp
                805                 810                 815

Asp Thr Phe Leu Pro Val Pro Glu Tyr Ile Asn Gln Ser Val Pro Lys
            820                 825                 830

Arg Pro Ala Gly Ser Val Gln Asn Pro Val Tyr His Asn Gln Pro Leu
        835                 840                 845

Asn Pro Ala Pro Ser Arg Asp Pro His Tyr Gln Asp Pro His Ser Thr
    850                 855                 860

Ala Val Gly Asn Pro Glu Tyr Leu Asn Thr Val Gln Pro Thr Cys Val
865                 870                 875                 880

Asn Ser Thr Phe Asp Ser Pro Ala His Trp Ala Gln Lys Gly Ser His
                885                 890                 895

Gln Ile Ser Leu Asp Asn Pro Asp Tyr Gln Gln Asp Phe Phe Pro Lys
            900                 905                 910

Glu Ala Lys Pro Asn Gly Ile Phe Lys Gly Ser Thr Ala Glu Asn Ala
        915                 920                 925

Glu Tyr Leu Arg Val Ala Pro Gln Ser Ser Glu Phe Ile Gly Ala
    930                 935                 940
```

<210> 9
<211> 4530
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (151)..(3918)
<223> HER-2 coding sequence

<400> 9
aattctcgag ctcgtcgacc ggtcgacgag ctcgagggtc gacgagctcg agggcgcgcg 60

cccggccccc acccctcgca gcaccccgcg ccccgcgccc tcccagccgg gtccagccgg 120

agccatgggg ccggagccgc agtgagcacc atg gag ctg gcg gcc ttg tgc cgc 174
Met Glu Leu Ala Ala Leu Cys Arg
1 5

tgg ggg ctc ctc ctc gcc ctc ttg ccc ccc gga gcc gcg agc acc caa 222
Trp Gly Leu Leu Leu Ala Leu Leu Pro Pro Gly Ala Ala Ser Thr Gln
10 15 20

gtg tgc acc ggc aca gac atg aag ctg cgg ctc cct gcc agt ccc gag 270
Val Cys Thr Gly Thr Asp Met Lys Leu Arg Leu Pro Ala Ser Pro Glu
25 30 35 40

acc cac ctg gac atg ctc cgc cac ctc tac cag ggc tgc cag gtg gtg 318
Thr His Leu Asp Met Leu Arg His Leu Tyr Gln Gly Cys Gln Val Val
45 50 55

cag gga aac ctg gaa ctc acc tac ctg ccc acc aat gcc agc ctg tcc 366
Gln Gly Asn Leu Glu Leu Thr Tyr Leu Pro Thr Asn Ala Ser Leu Ser
60 65 70

ttc ctg cag gat atc cag gag gtg cag ggc tac gtg ctc atc gct cac 414
Phe Leu Gln Asp Ile Gln Glu Val Gln Gly Tyr Val Leu Ile Ala His
75 80 85

aac caa gtg agg cag gtc cca ctg cag agg ctg cgg att gtg cga ggc 462
Asn Gln Val Arg Gln Val Pro Leu Gln Arg Leu Arg Ile Val Arg Gly
90 95 100

acc cag ctc ttt gag gac aac tat gcc ctg gcc gtg cta gac aat gga 510
Thr Gln Leu Phe Glu Asp Asn Tyr Ala Leu Ala Val Leu Asp Asn Gly
105 110 115 120

gac ccg ctg aac aat acc acc cct gtc aca ggg gcc tcc cca gga ggc 558
Asp Pro Leu Asn Asn Thr Thr Pro Val Thr Gly Ala Ser Pro Gly Gly
125 130 135

ctg cgg gag ctg cag ctt cga agc ctc aca gag atc ttg aaa gga ggg 606
Leu Arg Glu Leu Gln Leu Arg Ser Leu Thr Glu Ile Leu Lys Gly Gly
140 145 150

gtc ttg atc cag cgg aac ccc cag ctc tgc tac cag gac acg att ttg 654
Val Leu Ile Gln Arg Asn Pro Gln Leu Cys Tyr Gln Asp Thr Ile Leu
155 160 165

tgg aag gac atc ttc cac aag aac aac cag ctg gct ctc aca ctg ata 702
Trp Lys Asp Ile Phe His Lys Asn Asn Gln Leu Ala Leu Thr Leu Ile
170 175 180

```
gac acc aac cgc tct cgg gcc tgc cac ccc tgt tct ccg atg tgt aag      750
Asp Thr Asn Arg Ser Arg Ala Cys His Pro Cys Ser Pro Met Cys Lys
185                 190                 195                 200

ggc tcc cgc tgc tgg gga gag agt tct gag gat tgt cag agc ctg acg      798
Gly Ser Arg Cys Trp Gly Glu Ser Ser Glu Asp Cys Gln Ser Leu Thr
                205                 210                 215

cgc act gtc tgt gcc ggt ggc tgt gcc cgc tgc aag ggg cca ctg ccc      846
Arg Thr Val Cys Ala Gly Gly Cys Ala Arg Cys Lys Gly Pro Leu Pro
            220                 225                 230

act gac tgc tgc cat gag cag tgt gct gcc ggc tgc acg ggc ccc aag      894
Thr Asp Cys Cys His Glu Gln Cys Ala Ala Gly Cys Thr Gly Pro Lys
        235                 240                 245

cac tct gac tgc ctg gcc tgc ctc cac ttc aac cac agt ggc atc tgt      942
His Ser Asp Cys Leu Ala Cys Leu His Phe Asn His Ser Gly Ile Cys
    250                 255                 260

gag ctg cac tgc cca gcc ctg gtc acc tac aac aca gac acg ttt gag      990
Glu Leu His Cys Pro Ala Leu Val Thr Tyr Asn Thr Asp Thr Phe Glu
265                 270                 275                 280

tcc atg ccc aat ccc gag ggc cgg tat aca ttc ggc gcc agc tgt gtg     1038
Ser Met Pro Asn Pro Glu Gly Arg Tyr Thr Phe Gly Ala Ser Cys Val
                285                 290                 295

act gcc tgt ccc tac aac tac ctt tct acg gac gtg gga tcc tgc acc     1086
Thr Ala Cys Pro Tyr Asn Tyr Leu Ser Thr Asp Val Gly Ser Cys Thr
            300                 305                 310

ctc gtc tgc ccc ctg cac aac caa gag gtg aca gca gag gat gga aca     1134
Leu Val Cys Pro Leu His Asn Gln Glu Val Thr Ala Glu Asp Gly Thr
        315                 320                 325

cag cgg tgt gag aag tgc agc aag ccc tgt gcc cga gtg tgc tat ggt     1182
Gln Arg Cys Glu Lys Cys Ser Lys Pro Cys Ala Arg Val Cys Tyr Gly
    330                 335                 340

ctg ggc atg gag cac ttg cga gag gtg agg gca gtt acc agt gcc aat     1230
Leu Gly Met Glu His Leu Arg Glu Val Arg Ala Val Thr Ser Ala Asn
345                 350                 355                 360

atc cag gag ttt gct ggc tgc aag aag atc ttt ggg agc ctg gca ttt     1278
Ile Gln Glu Phe Ala Gly Cys Lys Lys Ile Phe Gly Ser Leu Ala Phe
                365                 370                 375

ctg ccg gag agc ttt gat ggg gac cca gcc tcc aac act gcc ccg ctc     1326
Leu Pro Glu Ser Phe Asp Gly Asp Pro Ala Ser Asn Thr Ala Pro Leu
            380                 385                 390

cag cca gag cag ctc caa gtg ttt gag act ctg gaa gag atc aca ggt     1374
Gln Pro Glu Gln Leu Gln Val Phe Glu Thr Leu Glu Glu Ile Thr Gly
        395                 400                 405

tac cta tac atc tca gca tgg ccg gac agc ctg cct gac ctc agc gtc     1422
Tyr Leu Tyr Ile Ser Ala Trp Pro Asp Ser Leu Pro Asp Leu Ser Val
    410                 415                 420

ttc cag aac ctg caa gta atc cgg gga cga att ctg cac aat ggc gcc     1470
Phe Gln Asn Leu Gln Val Ile Arg Gly Arg Ile Leu His Asn Gly Ala
425                 430                 435                 440
```

```
tac tcg ctg acc ctg caa ggg ctg ggc atc agc tgg ctg ggg ctg cgc      1518
Tyr Ser Leu Thr Leu Gln Gly Leu Gly Ile Ser Trp Leu Gly Leu Arg
                445                 450                 455

tca ctg agg gaa ctg ggc agt gga ctg gcc ctc atc cac cat aac acc      1566
Ser Leu Arg Glu Leu Gly Ser Gly Leu Ala Leu Ile His His Asn Thr
            460                 465                 470

cac ctc tgc ttc gtg cac acg gtg ccc tgg gac cag ctc ttt cgg aac      1614
His Leu Cys Phe Val His Thr Val Pro Trp Asp Gln Leu Phe Arg Asn
        475                 480                 485

ccg cac caa gct ctg ctc cac act gcc aac cgg cca gag gac gag tgt      1662
Pro His Gln Ala Leu Leu His Thr Ala Asn Arg Pro Glu Asp Glu Cys
    490                 495                 500

gtg ggc gag ggc ctg gcc tgc cac cag ctg tgc gcc cga ggg cac tgc      1710
Val Gly Glu Gly Leu Ala Cys His Gln Leu Cys Ala Arg Gly His Cys
505                 510                 515                 520

tgg ggt cca ggg ccc acc cag tgt gtc aac tgc agc cag ttc ctt cgg      1758
Trp Gly Pro Gly Pro Thr Gln Cys Val Asn Cys Ser Gln Phe Leu Arg
                525                 530                 535

ggc cag gag tgc gtg gag gaa tgc cga gta ctg cag ggg ctc ccc agg      1806
Gly Gln Glu Cys Val Glu Glu Cys Arg Val Leu Gln Gly Leu Pro Arg
            540                 545                 550

gag tat gtg aat gcc agg cac tgt ttg ccg tgc cac cct gag tgt cag      1854
Glu Tyr Val Asn Ala Arg His Cys Leu Pro Cys His Pro Glu Cys Gln
        555                 560                 565

ccc cag aat ggc tca gtg acc tgt ttt gga ccg gag gct gac cag tgt      1902
Pro Gln Asn Gly Ser Val Thr Cys Phe Gly Pro Glu Ala Asp Gln Cys
    570                 575                 580

gtg gcc tgt gcc cac tat aag gac cct ccc ttc tgc gtg gcc cgc tgc      1950
Val Ala Cys Ala His Tyr Lys Asp Pro Pro Phe Cys Val Ala Arg Cys
585                 590                 595                 600

ccc agc ggt gtg aaa cct gac ctc tcc tac atg ccc atc tgg aag ttt      1998
Pro Ser Gly Val Lys Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe
                605                 610                 615

cca gat gag gag ggc gca tgc cag cct tgc ccc atc aac tgc acc cac      2046
Pro Asp Glu Glu Gly Ala Cys Gln Pro Cys Pro Ile Asn Cys Thr His
            620                 625                 630

tcc tgt gtg gac ctg gat gac aag ggc tgc ccc gcc gag cag aga gcc      2094
Ser Cys Val Asp Leu Asp Asp Lys Gly Cys Pro Ala Glu Gln Arg Ala
        635                 640                 645

agc cct ctg acg tcc atc gtc tct gcg gtg gtt ggc att ctg ctg gtc      2142
Ser Pro Leu Thr Ser Ile Val Ser Ala Val Val Gly Ile Leu Leu Val
    650                 655                 660

gtg gtc ttg ggg gtg gtc ttt ggg atc ctc atc aag cga cgg cag cag      2190
Val Val Leu Gly Val Val Phe Gly Ile Leu Ile Lys Arg Arg Gln Gln
665                 670                 675                 680

aag atc cgg aag tac acg atg cgg aga ctg ctg cag gaa acg gag ctg      2238
Lys Ile Arg Lys Tyr Thr Met Arg Arg Leu Leu Gln Glu Thr Glu Leu
                685                 690                 695
```

```
gtg gag ccg ctg aca cct agc gga gcg atg ccc aac cag gcg cag atg     2286
Val Glu Pro Leu Thr Pro Ser Gly Ala Met Pro Asn Gln Ala Gln Met
            700                 705                 710

cgg atc ctg aaa gag acg gag ctg agg aag gtg aag gtg ctt gga tct     2334
Arg Ile Leu Lys Glu Thr Glu Leu Arg Lys Val Lys Val Leu Gly Ser
        715                 720                 725

ggc gct ttt ggc aca gtc tac aag ggc atc tgg atc cct gat ggg gag     2382
Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile Trp Ile Pro Asp Gly Glu
    730                 735                 740

aat gtg aaa att cca gtg gcc atc aaa gtg ttg agg gaa aac aca tcc     2430
Asn Val Lys Ile Pro Val Ala Ile Lys Val Leu Arg Glu Asn Thr Ser
745                 750                 755                 760

ccc aaa gcc aac aaa gaa atc tta gac gaa gca tac gtg atg gct ggt     2478
Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met Ala Gly
                765                 770                 775

gtg ggc tcc cca tat gtc tcc cgc ctt ctg ggc atc tgc ctg aca tcc     2526
Val Gly Ser Pro Tyr Val Ser Arg Leu Leu Gly Ile Cys Leu Thr Ser
            780                 785                 790

acg gtg cag ctg gtg aca cag ctt atg ccc tat ggc tgc ctc tta gac     2574
Thr Val Gln Leu Val Thr Gln Leu Met Pro Tyr Gly Cys Leu Leu Asp
        795                 800                 805

cat gtc cgg gaa aac cgc gga cgc ctg ggc tcc cag gac ctg ctg aac     2622
His Val Arg Glu Asn Arg Gly Arg Leu Gly Ser Gln Asp Leu Leu Asn
    810                 815                 820

tgg tgt atg cag att gcc aag ggg atg agc tac ctg gag gat gtg cgg     2670
Trp Cys Met Gln Ile Ala Lys Gly Met Ser Tyr Leu Glu Asp Val Arg
825                 830                 835                 840

ctc gta cac agg gac ttg gcc gct cgg aac gtg ctg gtc aag agt ccc     2718
Leu Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Ser Pro
                845                 850                 855

aac cat gtc aaa att aca gac ttc ggg ctg gct cgg ctg ctg gac att     2766
Asn His Val Lys Ile Thr Asp Phe Gly Leu Ala Arg Leu Leu Asp Ile
            860                 865                 870

gac gag aca gag tac cat gca gat ggg ggc aag gtg ccc atc aag tgg     2814
Asp Glu Thr Glu Tyr His Ala Asp Gly Gly Lys Val Pro Ile Lys Trp
        875                 880                 885

atg gcg ctg gag tcc att ctc cgc cgg cgg ttc acc cac cag agt gat     2862
Met Ala Leu Glu Ser Ile Leu Arg Arg Arg Phe Thr His Gln Ser Asp
    890                 895                 900

gtg tgg agt tat ggt gtg act gtg tgg gag ctg atg act ttt ggg gcc     2910
Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ala
905                 910                 915                 920

aaa cct tac gat ggg atc cca gcc cgg gag atc cct gac ctg ctg gaa     2958
Lys Pro Tyr Asp Gly Ile Pro Ala Arg Glu Ile Pro Asp Leu Leu Glu
                925                 930                 935

aag ggg gag cgg ctg ccc cag ccc ccc atc tgc acc att gat gtc tac     3006
Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr
            940                 945                 950
```

```
atg atc atg gtc aaa tgt tgg atg att gac tct gaa tgt cgg cca aga     3054
Met Ile Met Val Lys Cys Trp Met Ile Asp Ser Glu Cys Arg Pro Arg
        955                 960                 965

ttc cgg gag ttg gtg tct gaa ttc tcc cgc atg gcc agg gac ccc cag     3102
Phe Arg Glu Leu Val Ser Glu Phe Ser Arg Met Ala Arg Asp Pro Gln
    970                 975                 980

cgc ttt gtg gtc atc cag aat gag gac ttg ggc cca gcc agt ccc ttg     3150
Arg Phe Val Val Ile Gln Asn Glu Asp Leu Gly Pro Ala Ser Pro Leu
985                 990                 995                 1000

gac agc acc ttc tac  cgc tca ctg ctg gag  gac gat gac atg ggg       3195
Asp Ser Thr Phe Tyr  Arg Ser Leu Leu Glu  Asp Asp Asp Met Gly
                1005                1010                1015

gac ctg gtg gat gct  gag gag tat ctg gta  ccc cag cag ggc ttc       3240
Asp Leu Val Asp Ala  Glu Glu Tyr Leu Val  Pro Gln Gln Gly Phe
                1020                1025                1030

ttc tgt cca gac cct  gcc ccg ggc gct ggg  ggc atg gtc cac cac       3285
Phe Cys Pro Asp Pro  Ala Pro Gly Ala Gly  Gly Met Val His His
                1035                1040                1045

agg cac cgc agc tca  tct acc agg agt ggc  ggt ggg gac ctg aca       3330
Arg His Arg Ser Ser  Ser Thr Arg Ser Gly  Gly Gly Asp Leu Thr
                1050                1055                1060

cta ggg ctg gag ccc  tct gaa gag gag gcc  ccc agg tct cca ctg       3375
Leu Gly Leu Glu Pro  Ser Glu Glu Glu Ala  Pro Arg Ser Pro Leu
                1065                1070                1075

gca ccc tcc gaa ggg  gct ggc tcc gat gta  ttt gat ggt gac ctg       3420
Ala Pro Ser Glu Gly  Ala Gly Ser Asp Val  Phe Asp Gly Asp Leu
                1080                1085                1090

gga atg ggg gca gcc  aag ggg ctg caa agc  ctc ccc aca cat gac       3465
Gly Met Gly Ala Ala  Lys Gly Leu Gln Ser  Leu Pro Thr His Asp
                1095                1100                1105

ccc agc cct cta cag  cgg tac agt gag gac  ccc aca gta ccc ctg       3510
Pro Ser Pro Leu Gln  Arg Tyr Ser Glu Asp  Pro Thr Val Pro Leu
                1110                1115                1120

ccc tct gag act gat  ggc tac gtt gcc ccc  ctg acc tgc agc ccc       3555
Pro Ser Glu Thr Asp  Gly Tyr Val Ala Pro  Leu Thr Cys Ser Pro
                1125                1130                1135

cag cct gaa tat gtg  aac cag cca gat gtt  cgg ccc cag ccc cct       3600
Gln Pro Glu Tyr Val  Asn Gln Pro Asp Val  Arg Pro Gln Pro Pro
                1140                1145                1150

tcg ccc cga gag ggc  cct ctg cct gct gcc  cga cct gct ggt gcc       3645
Ser Pro Arg Glu Gly  Pro Leu Pro Ala Ala  Arg Pro Ala Gly Ala
                1155                1160                1165

act ctg gaa agg gcc  aag act ctc tcc cca  ggg aag aat ggg gtc       3690
Thr Leu Glu Arg Ala  Lys Thr Leu Ser Pro  Gly Lys Asn Gly Val
                1170                1175                1180

gtc aaa gac gtt ttt  gcc ttt ggg ggt gcc  gtg gag aac ccc gag       3735
Val Lys Asp Val Phe  Ala Phe Gly Gly Ala  Val Glu Asn Pro Glu
                1185                1190                1195
```

```
tac ttg aca ccc cag  gga gga gct gcc cct  cag ccc cac cct cct       3780
Tyr Leu Thr Pro Gln  Gly Gly Ala Ala Pro  Gln Pro His Pro Pro
                1200                1205                1210

cct gcc ttc agc cca  gcc ttc gac aac ctc  tat tac tgg gac cag       3825
Pro Ala Phe Ser Pro  Ala Phe Asp Asn Leu  Tyr Tyr Trp Asp Gln
                1215                1220                1225

gac cca cca gag cgg  ggg gct cca ccc agc  acc ttc aaa ggg aca       3870
Asp Pro Pro Glu Arg  Gly Ala Pro Pro Ser  Thr Phe Lys Gly Thr
                1230                1235                1240

cct acg gca gag aac  cca gag tac ctg ggt  ctg gac gtg cca gtg       3915
Pro Thr Ala Glu Asn  Pro Glu Tyr Leu Gly  Leu Asp Val Pro Val
                1245                1250                1255

tga accagaaggc caagtccgca gaagccctga tgtgtcctca gggagcaggg          3968

aaggcctgac ttctgctggc atcaagaggt gggagggccc tccgaccact tccaggggaa   4028

cctgccatgc caggaacctg tcctaaggaa ccttccttcc tgcttgagtt cccagatggc   4088

tggaaggggt ccagcctcgt tggaagagga acagcactgg ggagtctttg tggattctga   4148

ggccctgccc aatgagactc tagggtccag tggatgccac agcccagctt ggccctttcc   4208

ttccagatcc tgggtactga aagccttagg gaagctggcc tgagagggga agcggcccta   4268

agggagtgtc taagaacaaa agcgacccat tcagagactg tccctgaaac ctagtactgc   4328

cccccatgag gaaggaacag caatggtgtc agtatccagg ctttgtacag agtgcttttc   4388

tgtttagttt ttacttttt tgttttgttt ttttaaagac gaaataaaga cccaggggag   4448

aatgggtgtt gtatggggag gcaagtgtgg ggggtccttc tccacaccca ctttgtccat   4508

ttgcaaatat attttggaaa ac                                           4530


<210>  10
<211>  1255
<212>  PRT
<213>  Homo sapiens

<400>  10

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                  10                  15


Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30


Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45


Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60


Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80
```

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu
340 345 350

Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys
355 360 365

Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp
370 375 380

Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe
385 390 395 400

Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro
405 410 415

Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg
420 425 430

Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu
435 440 445

Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly
450 455 460

Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val His Thr Val
465 470 475 480

Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr
485 490 495

Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His
500 505 510

Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys
515 520 525

Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys
530 535 540

Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys
545 550 555 560

Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys
565 570 575

Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp
580 585 590

Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu
595 600 605

Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln
610 615 620

Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys
625 630 635 640

Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Ile Val Ser
645 650 655

Ala Val Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly
660 665 670

Ile Leu Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg
675 680 685

Arg Leu Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
690 695 700

Ala Met Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu
705 710 715 720

Arg Lys Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
725 730 735

Gly Ile Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile
740 745 750

Lys Val Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
755 760 765

Asp Glu Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg
770 775 780

Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu
785 790 795 800

Met Pro Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg
805 810 815

Leu Gly Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly
820 825 830

Met Ser Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala
835 840 845

Arg Asn Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe
850 855 860

Gly Leu Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp
865 870 875 880

Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu Arg
885 890 895

Arg Arg Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
900 905 910

Trp Glu Leu Met Thr Phe Gly Ala Lys Pro Tyr Asp Gly Ile Pro Ala
915 920 925

Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
930 935 940

Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
945 950 955 960

Ile Asp Ser Glu Cys Arg Pro Arg Phe Arg Glu Leu Val Ser Glu Phe
965 970 975

Ser Arg Met Ala Arg Asp Pro Gln Arg Phe Val Val Ile Gln Asn Glu
980 985 990

Asp Leu Gly Pro Ala Ser Pro Leu Asp Ser Thr Phe Tyr Arg Ser Leu
995 1000 1005

Leu Glu Asp Asp Asp Met Gly Asp Leu Val Asp Ala Glu Glu Tyr
1010 1015 1020

Leu Val Pro Gln Gln Gly Phe Phe Cys Pro Asp Pro Ala Pro Gly
1025 1030 1035

Ala Gly Gly Met Val His His Arg His Arg Ser Ser Ser Thr Arg
1040 1045 1050

Ser Gly Gly Gly Asp Leu Thr Leu Gly Leu Glu Pro Ser Glu Glu
1055 1060 1065

Glu Ala Pro Arg Ser Pro Leu Ala Pro Ser Glu Gly Ala Gly Ser
1070 1075 1080

Asp Val Phe Asp Gly Asp Leu Gly Met Gly Ala Ala Lys Gly Leu
1085 1090 1095

Gln Ser Leu Pro Thr His Asp Pro Ser Pro Leu Gln Arg Tyr Ser
1100 1105 1110

Glu Asp Pro Thr Val Pro Leu Pro Ser Glu Thr Asp Gly Tyr Val
1115 1120 1125

Ala Pro Leu Thr Cys Ser Pro Gln Pro Glu Tyr Val Asn Gln Pro
1130 1135 1140

Asp Val Arg Pro Gln Pro Pro Ser Pro Arg Glu Gly Pro Leu Pro
1145 1150 1155

Ala Ala Arg Pro Ala Gly Ala Thr Leu Glu Arg Ala Lys Thr Leu
1160 1165 1170

Ser Pro Gly Lys Asn Gly Val Val Lys Asp Val Phe Ala Phe Gly
1175 1180 1185

Gly Ala Val Glu Asn Pro Glu Tyr Leu Thr Pro Gln Gly Gly Ala
1190 1195 1200

Ala Pro Gln Pro His Pro Pro Pro Ala Phe Ser Pro Ala Phe Asp
1205 1210 1215

Asn Leu Tyr Tyr Trp Asp Gln Asp Pro Pro Glu Arg Gly Ala Pro
1220 1225 1230

Pro Ser Thr Phe Lys Gly Thr Pro Thr Ala Glu Asn Pro Glu Tyr
1235 1240 1245

Leu Gly Leu Asp Val Pro Val
1250 1255

<210> 11
<211> 4975
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (199)..(4227)

<400> 11
ctctcacaca cacacacccc tcccctgcca tccctccccg gactccggct ccggctccga 60

ttgcaatttg caacctccgc tgccgtcgcc gcagcagcca ccaattcgcc agcggttcag 120

gtggctcttg cctcgatgtc ctagcctagg ggcccccggg ccggacttgg ctgggctccc 180

ttcaccctct gcggagtc atg agg gcg aac gac gct ctg cag gtg ctg ggc 231

```
                    Met Arg Ala Asn Asp Ala Leu Gln Val Leu Gly
                    1               5                   10

ttg ctt ttc agc ctg gcc cgg ggc tcc gag gtg ggc aac tct cag gca      279
Leu Leu Phe Ser Leu Ala Arg Gly Ser Glu Val Gly Asn Ser Gln Ala
            15                  20                  25

gtg tgt cct ggg act ctg aat ggc ctg agt gtg acc ggc gat gct gag      327
Val Cys Pro Gly Thr Leu Asn Gly Leu Ser Val Thr Gly Asp Ala Glu
        30                  35                  40

aac caa tac cag aca ctg tac aag ctc tac gag agg tgt gag gtg gtg      375
Asn Gln Tyr Gln Thr Leu Tyr Lys Leu Tyr Glu Arg Cys Glu Val Val
    45                  50                  55

atg ggg aac ctt gag att gtg ctc acg gga cac aat gcc gac ctc tcc      423
Met Gly Asn Leu Glu Ile Val Leu Thr Gly His Asn Ala Asp Leu Ser
60                  65                  70                  75

ttc ctg cag tgg att cga gaa gtg aca ggc tat gtc ctc gtg gcc atg      471
Phe Leu Gln Trp Ile Arg Glu Val Thr Gly Tyr Val Leu Val Ala Met
                80                  85                  90

aat gaa ttc tct act cta cca ttg ccc aac ctc cgc gtg gtg cga ggg      519
Asn Glu Phe Ser Thr Leu Pro Leu Pro Asn Leu Arg Val Val Arg Gly
            95                  100                 105

acc cag gtc tac gat ggg aag ttt gcc atc ttc gtc atg ttg aac tat      567
Thr Gln Val Tyr Asp Gly Lys Phe Ala Ile Phe Val Met Leu Asn Tyr
        110                 115                 120

aac acc aac tcc agc cac gct ctg cgc cag ctc cgc ttg act cag ctc      615
Asn Thr Asn Ser Ser His Ala Leu Arg Gln Leu Arg Leu Thr Gln Leu
    125                 130                 135

acc gag att ctg tca ggg ggt gtt tat att gag aag aac gat aag ctt      663
Thr Glu Ile Leu Ser Gly Gly Val Tyr Ile Glu Lys Asn Asp Lys Leu
140                 145                 150                 155

tgt cac atg gac aca att gac tgg agg gac atc gtg agg gac cga gat      711
Cys His Met Asp Thr Ile Asp Trp Arg Asp Ile Val Arg Asp Arg Asp
                160                 165                 170

gct gag ata gtg gtg aag gac aat ggc aga agc tgt ccc ccc tgt cat      759
Ala Glu Ile Val Val Lys Asp Asn Gly Arg Ser Cys Pro Pro Cys His
            175                 180                 185

gag gtt tgc aag ggg cga tgc tgg ggt cct gga tca gaa gac tgc cag      807
Glu Val Cys Lys Gly Arg Cys Trp Gly Pro Gly Ser Glu Asp Cys Gln
        190                 195                 200

aca ttg acc aag acc atc tgt gct cct cag tgt aat ggt cac tgc ttt      855
Thr Leu Thr Lys Thr Ile Cys Ala Pro Gln Cys Asn Gly His Cys Phe
    205                 210                 215

ggg ccc aac ccc aac cag tgc tgc cat gat gag tgt gcc ggg ggc tgc      903
Gly Pro Asn Pro Asn Gln Cys Cys His Asp Glu Cys Ala Gly Gly Cys
220                 225                 230                 235

tca ggc cct cag gac aca gac tgc ttt gcc tgc cgg cac ttc aat gac      951
Ser Gly Pro Gln Asp Thr Asp Cys Phe Ala Cys Arg His Phe Asn Asp
                240                 245                 250

agt gga gcc tgt gta cct cgc tgt cca cag cct ctt gtc tac aac aag      999
```

```
Ser Gly Ala Cys Val Pro Arg Cys Pro Gln Pro Leu Val Tyr Asn Lys
            255                 260                 265

cta act ttc cag ctg gaa ccc aat ccc cac acc aag tat cag tat gga     1047
Leu Thr Phe Gln Leu Glu Pro Asn Pro His Thr Lys Tyr Gln Tyr Gly
        270                 275                 280

gga gtt tgt gta gcc agc tgt ccc cat aac ttt gtg gtg gat caa aca     1095
Gly Val Cys Val Ala Ser Cys Pro His Asn Phe Val Val Asp Gln Thr
    285                 290                 295

tcc tgt gtc agg gcc tgt cct cct gac aag atg gaa gta gat aaa aat     1143
Ser Cys Val Arg Ala Cys Pro Pro Asp Lys Met Glu Val Asp Lys Asn
300                 305                 310                 315

ggg ctc aag atg tgt gag cct tgt ggg gga cta tgt ccc aaa gcc tgt     1191
Gly Leu Lys Met Cys Glu Pro Cys Gly Gly Leu Cys Pro Lys Ala Cys
                320                 325                 330

gag gga aca ggc tct ggg agc cgc ttc cag act gtg gac tcg agc aac     1239
Glu Gly Thr Gly Ser Gly Ser Arg Phe Gln Thr Val Asp Ser Ser Asn
            335                 340                 345

att gat gga ttt gtg aac tgc acc aag atc ctg ggc aac ctg gac ttt     1287
Ile Asp Gly Phe Val Asn Cys Thr Lys Ile Leu Gly Asn Leu Asp Phe
        350                 355                 360

ctg atc acc ggc ctc aat gga gac ccc tgg cac aag atc cct gcc ctg     1335
Leu Ile Thr Gly Leu Asn Gly Asp Pro Trp His Lys Ile Pro Ala Leu
    365                 370                 375

gac cca gag aag ctc aat gtc ttc cgg aca gta cgg gag atc aca ggt     1383
Asp Pro Glu Lys Leu Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly
380                 385                 390                 395

tac ctg aac atc cag tcc tgg ccg ccc cac atg cac aac ttc agt gtt     1431
Tyr Leu Asn Ile Gln Ser Trp Pro Pro His Met His Asn Phe Ser Val
                400                 405                 410

ttt tcc aat ttg aca acc att gga ggc aga agc ctc tac aac cgg ggc     1479
Phe Ser Asn Leu Thr Thr Ile Gly Gly Arg Ser Leu Tyr Asn Arg Gly
            415                 420                 425

ttc tca ttg ttg atc atg aag aac ttg aat gtc aca tct ctg ggc ttc     1527
Phe Ser Leu Leu Ile Met Lys Asn Leu Asn Val Thr Ser Leu Gly Phe
        430                 435                 440

cga tcc ctg aag gaa att agt gct ggg cgt atc tat ata agt gcc aat     1575
Arg Ser Leu Lys Glu Ile Ser Ala Gly Arg Ile Tyr Ile Ser Ala Asn
    445                 450                 455

agg cag ctc tgc tac cac cac tct ttg aac tgg acc aag gtg ctt cgg     1623
Arg Gln Leu Cys Tyr His His Ser Leu Asn Trp Thr Lys Val Leu Arg
460                 465                 470                 475

ggg cct acg gaa gag cga cta gac atc aag cat aat cgg ccg cgc aga     1671
Gly Pro Thr Glu Glu Arg Leu Asp Ile Lys His Asn Arg Pro Arg Arg
                480                 485                 490

gac tgc gtg gca gag ggc aaa gtg tgt gac cca ctg tgc tcc tct ggg     1719
Asp Cys Val Ala Glu Gly Lys Val Cys Asp Pro Leu Cys Ser Ser Gly
            495                 500                 505

gga tgc tgg ggc cca ggc cct ggt cag tgc ttg tcc tgt cga aat tat     1767
```

```
Gly Cys Trp Gly Pro Gly Pro Gly Gln Cys Leu Ser Cys Arg Asn Tyr
        510                 515                 520

agc cga gga ggt gtc tgt gtg acc cac tgc aac ttt ctg aat ggg gag     1815
Ser Arg Gly Gly Val Cys Val Thr His Cys Asn Phe Leu Asn Gly Glu
    525                 530                 535

cct cga gaa ttt gcc cat gag gcc gaa tgc ttc tcc tgc cac ccg gaa     1863
Pro Arg Glu Phe Ala His Glu Ala Glu Cys Phe Ser Cys His Pro Glu
540                 545                 550                 555

tgc caa ccc atg ggg ggc act gcc aca tgc aat ggc tcg ggc tct gat     1911
Cys Gln Pro Met Gly Gly Thr Ala Thr Cys Asn Gly Ser Gly Ser Asp
                560                 565                 570

act tgt gct caa tgt gcc cat ttt cga gat ggg ccc cac tgt gtg agc     1959
Thr Cys Ala Gln Cys Ala His Phe Arg Asp Gly Pro His Cys Val Ser
            575                 580                 585

agc tgc ccc cat gga gtc cta ggt gcc aag ggc cca atc tac aag tac     2007
Ser Cys Pro His Gly Val Leu Gly Ala Lys Gly Pro Ile Tyr Lys Tyr
        590                 595                 600

cca gat gtt cag aat gaa tgt cgg ccc tgc cat gag aac tgc acc cag     2055
Pro Asp Val Gln Asn Glu Cys Arg Pro Cys His Glu Asn Cys Thr Gln
    605                 610                 615

ggg tgt aaa gga cca gag ctt caa gac tgt tta gga caa aca ctg gtg     2103
Gly Cys Lys Gly Pro Glu Leu Gln Asp Cys Leu Gly Gln Thr Leu Val
620                 625                 630                 635

ctg atc ggc aaa acc cat ctg aca atg gct ttg aca gtg ata gca gga     2151
Leu Ile Gly Lys Thr His Leu Thr Met Ala Leu Thr Val Ile Ala Gly
                640                 645                 650

ttg gta gtg att ttc atg atg ctg ggc ggc act ttt ctc tac tgg cgt     2199
Leu Val Val Ile Phe Met Met Leu Gly Gly Thr Phe Leu Tyr Trp Arg
            655                 660                 665

ggg cgc cgg att cag aat aaa agg gct atg agg cga tac ttg gaa cgg     2247
Gly Arg Arg Ile Gln Asn Lys Arg Ala Met Arg Arg Tyr Leu Glu Arg
        670                 675                 680

ggt gag agc ata gag cct ctg gac ccc agt gag aag gct aac aaa gtc     2295
Gly Glu Ser Ile Glu Pro Leu Asp Pro Ser Glu Lys Ala Asn Lys Val
    685                 690                 695

ttg gcc aga atc ttc aaa gag aca gag cta agg aag ctt aaa gtg ctt     2343
Leu Ala Arg Ile Phe Lys Glu Thr Glu Leu Arg Lys Leu Lys Val Leu
700                 705                 710                 715

ggc tcg ggt gtc ttt gga act gtg cac aaa gga gtg tgg atc cct gag     2391
Gly Ser Gly Val Phe Gly Thr Val His Lys Gly Val Trp Ile Pro Glu
                720                 725                 730

ggt gaa tca atc aag att cca gtc tgc att aaa gtc att gag gac aag     2439
Gly Glu Ser Ile Lys Ile Pro Val Cys Ile Lys Val Ile Glu Asp Lys
            735                 740                 745

agt gga cgg cag agt ttt caa gct gtg aca gat cat atg ctg gcc att     2487
Ser Gly Arg Gln Ser Phe Gln Ala Val Thr Asp His Met Leu Ala Ile
        750                 755                 760

ggc agc ctg gac cat gcc cac att gta agg ctg ctg gga cta tgc cca     2535
```

```
Gly Ser Leu Asp His Ala His Ile Val Arg Leu Leu Gly Leu Cys Pro
    765                 770                 775

ggg tca tct ctg cag ctt gtc act caa tat ttg cct ctg ggt tct ctg     2583
Gly Ser Ser Leu Gln Leu Val Thr Gln Tyr Leu Pro Leu Gly Ser Leu
780                 785                 790                 795

ctg gat cat gtg aga caa cac cgg ggg gca ctg ggg cca cag ctg ctg     2631
Leu Asp His Val Arg Gln His Arg Gly Ala Leu Gly Pro Gln Leu Leu
                800                 805                 810

ctc aac tgg gga gta caa att gcc aag gga atg tac tac ctt gag gaa     2679
Leu Asn Trp Gly Val Gln Ile Ala Lys Gly Met Tyr Tyr Leu Glu Glu
            815                 820                 825

cat ggt atg gtg cat aga aac ctg gct gcc cga aac gtg cta ctc aag     2727
His Gly Met Val His Arg Asn Leu Ala Ala Arg Asn Val Leu Leu Lys
        830                 835                 840

tca ccc agt cag gtt cag gtg gca gat ttt ggt gtg gct gac ctg ctg     2775
Ser Pro Ser Gln Val Gln Val Ala Asp Phe Gly Val Ala Asp Leu Leu
    845                 850                 855

cct cct gat gat aag cag ctg cta tac agt gag gcc aag act cca att     2823
Pro Pro Asp Asp Lys Gln Leu Leu Tyr Ser Glu Ala Lys Thr Pro Ile
860                 865                 870                 875

aag tgg atg gcc ctt gag agt atc cac ttt ggg aaa tac aca cac cag     2871
Lys Trp Met Ala Leu Glu Ser Ile His Phe Gly Lys Tyr Thr His Gln
                880                 885                 890

agt gat gtc tgg agc tat ggt gtg aca gtt tgg gag ttg atg acc ttc     2919
Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe
            895                 900                 905

ggg gca gag ccc tat gca ggg cta cga ttg gct gaa gta cca gac ctg     2967
Gly Ala Glu Pro Tyr Ala Gly Leu Arg Leu Ala Glu Val Pro Asp Leu
        910                 915                 920

cta gag aag ggg gag cgg ttg gca cag ccc cag atc tgc aca att gat     3015
Leu Glu Lys Gly Glu Arg Leu Ala Gln Pro Gln Ile Cys Thr Ile Asp
    925                 930                 935

gtc tac atg gtg atg gtc aag tgt tgg atg att gat gag aac att cgc     3063
Val Tyr Met Val Met Val Lys Cys Trp Met Ile Asp Glu Asn Ile Arg
940                 945                 950                 955

cca acc ttt aaa gaa cta gcc aat gag ttc acc agg atg gcc cga gac     3111
Pro Thr Phe Lys Glu Leu Ala Asn Glu Phe Thr Arg Met Ala Arg Asp
                960                 965                 970

cca cca cgg tat ctg gtc ata aag aga gag agt ggg cct gga ata gcc     3159
Pro Pro Arg Tyr Leu Val Ile Lys Arg Glu Ser Gly Pro Gly Ile Ala
            975                 980                 985

cct ggg cca gag ccc cat ggt ctg aca aac aag aag cta  gag gaa gta     3207
Pro Gly Pro Glu Pro His Gly Leu Thr Asn Lys Lys Leu  Glu Glu Val
        990                 995                 1000

gag ctg  gag cca gaa cta gac  cta gac cta gac ttg  gaa gca gag       3252
Glu Leu  Glu Pro Glu Leu Asp  Leu Asp Leu Asp Leu  Glu Ala Glu
    1005                1010                1015

gag gac  aac ctg gca acc acc  aca ctg ggc tcc gcc  ctc agc cta       3297
```

```
Glu Asp  Asn Leu Ala Thr Thr  Thr Leu Gly Ser Ala  Leu Ser Leu
    1020                 1025                 1030

cca gtt  gga aca ctt aat cgg  cca cgt ggg agc cag  agc ctt tta      3342
Pro Val  Gly Thr Leu Asn Arg  Pro Arg Gly Ser Gln  Ser Leu Leu
    1035                 1040                 1045

agt cca  tca tct gga tac atg  ccc atg aac cag ggt  aat ctt ggg      3387
Ser Pro  Ser Ser Gly Tyr Met  Pro Met Asn Gln Gly  Asn Leu Gly
    1050                 1055                 1060

ggg tct  tgc cag gag tct gca  gtt tct ggg agc agt  gaa cgg tgc      3432
Gly Ser  Cys Gln Glu Ser Ala  Val Ser Gly Ser Ser  Glu Arg Cys
    1065                 1070                 1075

ccc cgt  cca gtc tct cta cac  cca atg cca cgg gga  tgc ctg gca      3477
Pro Arg  Pro Val Ser Leu His  Pro Met Pro Arg Gly  Cys Leu Ala
    1080                 1085                 1090

tca gag  tca tca gag ggg cat  gta aca ggc tct gag  gct gag ctc      3522
Ser Glu  Ser Ser Glu Gly His  Val Thr Gly Ser Glu  Ala Glu Leu
    1095                 1100                 1105

cag gag  aaa gtg tca atg tgt  aga agc cgg agc agg  agc cgg agc      3567
Gln Glu  Lys Val Ser Met Cys  Arg Ser Arg Ser Arg  Ser Arg Ser
    1110                 1115                 1120

cca cgg  cca cgc gga gat agc  gcc tac cat tcc cag  cgc cac agt      3612
Pro Arg  Pro Arg Gly Asp Ser  Ala Tyr His Ser Gln  Arg His Ser
    1125                 1130                 1135

ctg ctg  act cct gtt acc cca  ctc tcc cca ccc ggg  tta gag gaa      3657
Leu Leu  Thr Pro Val Thr Pro  Leu Ser Pro Pro Gly  Leu Glu Glu
    1140                 1145                 1150

gag gat  gtc aac ggt tat gtc  atg cca gat aca cac  ctc aaa ggt      3702
Glu Asp  Val Asn Gly Tyr Val  Met Pro Asp Thr His  Leu Lys Gly
    1155                 1160                 1165

act ccc  tcc tcc cgg gaa ggc  acc ctt tct tca gtg  ggt ctc agt      3747
Thr Pro  Ser Ser Arg Glu Gly  Thr Leu Ser Ser Val  Gly Leu Ser
    1170                 1175                 1180

tct gtc  ctg ggt act gaa gaa  gaa gat gaa gat gag  gag tat gaa      3792
Ser Val  Leu Gly Thr Glu Glu  Glu Asp Glu Asp Glu  Glu Tyr Glu
    1185                 1190                 1195

tac atg  aac cgg agg aga agg  cac agt cca cct cat  ccc cct agg      3837
Tyr Met  Asn Arg Arg Arg Arg  His Ser Pro Pro His  Pro Pro Arg
    1200                 1205                 1210

cca agt  tcc ctt gag gag ctg  ggt tat gag tac atg  gat gtg ggg      3882
Pro Ser  Ser Leu Glu Glu Leu  Gly Tyr Glu Tyr Met  Asp Val Gly
    1215                 1220                 1225

tca gac  ctc agt gcc tct ctg  ggc agc aca cag agt  tgc cca ctc      3927
Ser Asp  Leu Ser Ala Ser Leu  Gly Ser Thr Gln Ser  Cys Pro Leu
    1230                 1235                 1240

cac cct  gta ccc atc atg ccc  act gca ggc aca act  cca gat gaa      3972
His Pro  Val Pro Ile Met Pro  Thr Ala Gly Thr Thr  Pro Asp Glu
    1245                 1250                 1255

gac tat  gaa tat atg aat cgg  caa cga gat gga ggt  ggt cct ggg      4017
```

```
Asp Tyr  Glu Tyr Met Asn Arg  Gln Arg Asp Gly Gly  Gly Pro Gly
    1260                 1265                 1270

ggt gat  tat gca gcc atg ggg  gcc tgc cca gca tct  gag caa ggg      4062
Gly Asp  Tyr Ala Ala Met Gly  Ala Cys Pro Ala Ser  Glu Gln Gly
    1275                 1280                 1285

tat gaa  gag atg aga gct ttt  cag ggg cct gga cat  cag gcc ccc      4107
Tyr Glu  Glu Met Arg Ala Phe  Gln Gly Pro Gly His  Gln Ala Pro
    1290                 1295                 1300

cat gtc  cat tat gcc cgc cta  aaa act cta cgt agc  tta gag gct      4152
His Val  His Tyr Ala Arg Leu  Lys Thr Leu Arg Ser  Leu Glu Ala
    1305                 1310                 1315

aca gac  tct gcc ttt gat aac  cct gat tac tgg cat  agc agg ctt      4197
Thr Asp  Ser Ala Phe Asp Asn  Pro Asp Tyr Trp His  Ser Arg Leu
    1320                 1325                 1330

ttc ccc  aag gct aat gcc cag  aga acg taa ctcctgctcc ctgtggcact     4247
Phe Pro  Lys Ala Asn Ala Gln  Arg Thr
    1335                 1340

cagggagcat ttaatggcag ctagtgcctt tagagggtac cgtcttctcc ctattccctc   4307

tctctcccag gtcccagccc cttttcccca gtcccagaca attccattca atctttggag   4367

gcttttaaac attttgacac aaaattctta tggtatgtag ccagctgtgc actttcttct   4427

ctttcccaac cccaggaaag gttttcctta ttttgtgtgc tttcccagtc ccattcctca   4487

gcttcttcac aggcactcct ggagatatga aggattactc tccatatccc ttcctctcag   4547

gctcttgact acttggaact aggctcttat gtgtgccttt gtttcccatc agactgtcaa   4607

gaagaggaaa gggaggaaac ctagcagagg aaagtgtaat tttggtttat gactcttaac   4667

cccctagaaa gacagaagct taaaatctgt gaagaaagag gttaggagta gatattgatt   4727

actatcataa ttcagcactt aactatgagc caggcatcat actaaacttc acctacatta   4787

tctcacttag tcctttatca tccttaaaac aattctgtga catacatatt atctcatttt   4847

acacaaaggg aagtcgggca tggtggctca tgcctgtaat ctcagcactt tgggaggctg   4907

aggcagaagg attacctgag gcaaggagtt tgagaccagc ttagccaaca tagtaagacc   4967

cccatctc                                                           4975
```

<210> 12
<211> 1342
<212> PRT
<213> Homo sapiens

<400> 12

```
Met Arg Ala Asn Asp Ala Leu Gln Val Leu Gly Leu Leu Phe Ser Leu
1               5                  10                  15

Ala Arg Gly Ser Glu Val Gly Asn Ser Gln Ala Val Cys Pro Gly Thr
            20                  25                  30
```

Leu Asn Gly Leu Ser Val Thr Gly Asp Ala Glu Asn Gln Tyr Gln Thr
35 40 45

Leu Tyr Lys Leu Tyr Glu Arg Cys Glu Val Val Met Gly Asn Leu Glu
50 55 60

Ile Val Leu Thr Gly His Asn Ala Asp Leu Ser Phe Leu Gln Trp Ile
65 70 75 80

Arg Glu Val Thr Gly Tyr Val Leu Val Ala Met Asn Glu Phe Ser Thr
85 90 95

Leu Pro Leu Pro Asn Leu Arg Val Val Arg Gly Thr Gln Val Tyr Asp
100 105 110

Gly Lys Phe Ala Ile Phe Val Met Leu Asn Tyr Asn Thr Asn Ser Ser
115 120 125

His Ala Leu Arg Gln Leu Arg Leu Thr Gln Leu Thr Glu Ile Leu Ser
130 135 140

Gly Gly Val Tyr Ile Glu Lys Asn Asp Lys Leu Cys His Met Asp Thr
145 150 155 160

Ile Asp Trp Arg Asp Ile Val Arg Asp Arg Asp Ala Glu Ile Val Val
165 170 175

Lys Asp Asn Gly Arg Ser Cys Pro Pro Cys His Glu Val Cys Lys Gly
180 185 190

Arg Cys Trp Gly Pro Gly Ser Glu Asp Cys Gln Thr Leu Thr Lys Thr
195 200 205

Ile Cys Ala Pro Gln Cys Asn Gly His Cys Phe Gly Pro Asn Pro Asn
210 215 220

Gln Cys Cys His Asp Glu Cys Ala Gly Gly Cys Ser Gly Pro Gln Asp
225 230 235 240

Thr Asp Cys Phe Ala Cys Arg His Phe Asn Asp Ser Gly Ala Cys Val
245 250 255

Pro Arg Cys Pro Gln Pro Leu Val Tyr Asn Lys Leu Thr Phe Gln Leu
260 265 270

Glu Pro Asn Pro His Thr Lys Tyr Gln Tyr Gly Gly Val Cys Val Ala
275 280 285

```
Ser Cys Pro His Asn Phe Val Val Asp Gln Thr Ser Cys Val Arg Ala
    290                 295                 300

Cys Pro Pro Asp Lys Met Glu Val Asp Lys Asn Gly Leu Lys Met Cys
305                 310                 315                 320

Glu Pro Cys Gly Gly Leu Cys Pro Lys Ala Cys Glu Gly Thr Gly Ser
                325                 330                 335

Gly Ser Arg Phe Gln Thr Val Asp Ser Ser Asn Ile Asp Gly Phe Val
            340                 345                 350

Asn Cys Thr Lys Ile Leu Gly Asn Leu Asp Phe Leu Ile Thr Gly Leu
        355                 360                 365

Asn Gly Asp Pro Trp His Lys Ile Pro Ala Leu Asp Pro Glu Lys Leu
    370                 375                 380

Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly Tyr Leu Asn Ile Gln
385                 390                 395                 400

Ser Trp Pro Pro His Met His Asn Phe Ser Val Phe Ser Asn Leu Thr
                405                 410                 415

Thr Ile Gly Gly Arg Ser Leu Tyr Asn Arg Gly Phe Ser Leu Leu Ile
            420                 425                 430

Met Lys Asn Leu Asn Val Thr Ser Leu Gly Phe Arg Ser Leu Lys Glu
        435                 440                 445

Ile Ser Ala Gly Arg Ile Tyr Ile Ser Ala Asn Arg Gln Leu Cys Tyr
    450                 455                 460

His His Ser Leu Asn Trp Thr Lys Val Leu Arg Gly Pro Thr Glu Glu
465                 470                 475                 480

Arg Leu Asp Ile Lys His Asn Arg Pro Arg Arg Asp Cys Val Ala Glu
                485                 490                 495

Gly Lys Val Cys Asp Pro Leu Cys Ser Ser Gly Gly Cys Trp Gly Pro
            500                 505                 510

Gly Pro Gly Gln Cys Leu Ser Cys Arg Asn Tyr Ser Arg Gly Gly Val
        515                 520                 525

Cys Val Thr His Cys Asn Phe Leu Asn Gly Glu Pro Arg Glu Phe Ala
    530                 535                 540
```

```
His Glu Ala Glu Cys Phe Ser Cys His Pro Glu Cys Gln Pro Met Gly
545                 550                 555                 560

Gly Thr Ala Thr Cys Asn Gly Ser Gly Ser Asp Thr Cys Ala Gln Cys
                565                 570                 575

Ala His Phe Arg Asp Gly Pro His Cys Val Ser Ser Cys Pro His Gly
            580                 585                 590

Val Leu Gly Ala Lys Gly Pro Ile Tyr Lys Tyr Pro Asp Val Gln Asn
        595                 600                 605

Glu Cys Arg Pro Cys His Glu Asn Cys Thr Gln Gly Cys Lys Gly Pro
    610                 615                 620

Glu Leu Gln Asp Cys Leu Gly Gln Thr Leu Val Leu Ile Gly Lys Thr
625                 630                 635                 640

His Leu Thr Met Ala Leu Thr Val Ile Ala Gly Leu Val Val Ile Phe
                645                 650                 655

Met Met Leu Gly Gly Thr Phe Leu Tyr Trp Arg Gly Arg Arg Ile Gln
            660                 665                 670

Asn Lys Arg Ala Met Arg Arg Tyr Leu Glu Arg Gly Glu Ser Ile Glu
        675                 680                 685

Pro Leu Asp Pro Ser Glu Lys Ala Asn Lys Val Leu Ala Arg Ile Phe
    690                 695                 700

Lys Glu Thr Glu Leu Arg Lys Leu Lys Val Leu Gly Ser Gly Val Phe
705                 710                 715                 720

Gly Thr Val His Lys Gly Val Trp Ile Pro Glu Gly Glu Ser Ile Lys
                725                 730                 735

Ile Pro Val Cys Ile Lys Val Ile Glu Asp Lys Ser Gly Arg Gln Ser
            740                 745                 750

Phe Gln Ala Val Thr Asp His Met Leu Ala Ile Gly Ser Leu Asp His
        755                 760                 765

Ala His Ile Val Arg Leu Leu Gly Leu Cys Pro Gly Ser Ser Leu Gln
    770                 775                 780

Leu Val Thr Gln Tyr Leu Pro Leu Gly Ser Leu Leu Asp His Val Arg
785                 790                 795                 800
```

```
Gln His Arg Gly Ala Leu Gly Pro Gln Leu Leu Leu Asn Trp Gly Val
                805                 810                 815

Gln Ile Ala Lys Gly Met Tyr Tyr Leu Glu Glu His Gly Met Val His
            820                 825                 830

Arg Asn Leu Ala Ala Arg Asn Val Leu Leu Lys Ser Pro Ser Gln Val
        835                 840                 845

Gln Val Ala Asp Phe Gly Val Ala Asp Leu Leu Pro Pro Asp Asp Lys
    850                 855                 860

Gln Leu Leu Tyr Ser Glu Ala Lys Thr Pro Ile Lys Trp Met Ala Leu
865                 870                 875                 880

Glu Ser Ile His Phe Gly Lys Tyr Thr His Gln Ser Asp Val Trp Ser
                885                 890                 895

Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ala Glu Pro Tyr
            900                 905                 910

Ala Gly Leu Arg Leu Ala Glu Val Pro Asp Leu Leu Glu Lys Gly Glu
        915                 920                 925

Arg Leu Ala Gln Pro Gln Ile Cys Thr Ile Asp Val Tyr Met Val Met
    930                 935                 940

Val Lys Cys Trp Met Ile Asp Glu Asn Ile Arg Pro Thr Phe Lys Glu
945                 950                 955                 960

Leu Ala Asn Glu Phe Thr Arg Met Ala Arg Asp Pro Pro Arg Tyr Leu
                965                 970                 975

Val Ile Lys Arg Glu Ser Gly Pro Gly Ile Ala Pro Gly Pro Glu Pro
            980                 985                 990

His Gly Leu Thr Asn Lys Lys Leu  Glu Glu Val Glu Leu  Glu Pro Glu
        995                 1000                1005

Leu Asp  Leu Asp Leu Asp Leu  Glu Ala Glu Glu Asp  Asn Leu Ala
    1010                1015                1020

Thr Thr  Thr Leu Gly Ser Ala  Leu Ser Leu Pro Val  Gly Thr Leu
    1025                1030                1035

Asn Arg  Pro Arg Gly Ser Gln  Ser Leu Leu Ser Pro  Ser Ser Gly
    1040                1045                1050
```

Tyr Met Pro Met Asn Gln Gly Asn Leu Gly Gly Ser Cys Gln Glu
1055 1060 1065

Ser Ala Val Ser Gly Ser Ser Glu Arg Cys Pro Arg Pro Val Ser
1070 1075 1080

Leu His Pro Met Pro Arg Gly Cys Leu Ala Ser Glu Ser Ser Glu
1085 1090 1095

Gly His Val Thr Gly Ser Glu Ala Glu Leu Gln Glu Lys Val Ser
1100 1105 1110

Met Cys Arg Ser Arg Ser Arg Ser Arg Ser Pro Arg Pro Arg Gly
1115 1120 1125

Asp Ser Ala Tyr His Ser Gln Arg His Ser Leu Leu Thr Pro Val
1130 1135 1140

Thr Pro Leu Ser Pro Pro Gly Leu Glu Glu Glu Asp Val Asn Gly
1145 1150 1155

Tyr Val Met Pro Asp Thr His Leu Lys Gly Thr Pro Ser Ser Arg
1160 1165 1170

Glu Gly Thr Leu Ser Ser Val Gly Leu Ser Ser Val Leu Gly Thr
1175 1180 1185

Glu Glu Glu Asp Glu Asp Glu Glu Tyr Glu Tyr Met Asn Arg Arg
1190 1195 1200

Arg Arg His Ser Pro Pro His Pro Pro Arg Pro Ser Ser Leu Glu
1205 1210 1215

Glu Leu Gly Tyr Glu Tyr Met Asp Val Gly Ser Asp Leu Ser Ala
1220 1225 1230

Ser Leu Gly Ser Thr Gln Ser Cys Pro Leu His Pro Val Pro Ile
1235 1240 1245

Met Pro Thr Ala Gly Thr Thr Pro Asp Glu Asp Tyr Glu Tyr Met
1250 1255 1260

Asn Arg Gln Arg Asp Gly Gly Gly Pro Gly Gly Asp Tyr Ala Ala
1265 1270 1275

Met Gly Ala Cys Pro Ala Ser Glu Gln Gly Tyr Glu Glu Met Arg
1280 1285 1290

```
Ala Phe  Gln Gly Pro Gly His  Gln Ala Pro His Val  His Tyr Ala
    1295                 1300                 1305

Arg Leu  Lys Thr Leu Arg Ser  Leu Glu Ala Thr Asp  Ser Ala Phe
    1310                 1315                 1320

Asp Asn  Pro Asp Tyr Trp His  Ser Arg Leu Phe Pro  Lys Ala Asn
    1325                 1330                 1335

Ala Gln  Arg Thr
    1340
```

<210> 13
<211> 4975
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (199)..(4227)
<223> HER-3 mutant coding sequence

<220>
<221> mutation
<222> (1877)..(1877)
<223> mutation, comprising substitution of "a" instead of "g" at this position

<400> 13

```
ctctcacaca cacacacccc tcccctgcca tccctccccg gactccggct ccggctccga     60

ttgcaatttg caacctccgc tgccgtcgcc gcagcagcca ccaattcgcc agcggttcag    120

gtggctcttg cctcgatgtc ctagcctagg ggcccccggg ccggacttgg ctgggctccc    180

ttcaccctct gcggagtc atg agg gcg aac gac gct ctg cag gtg ctg ggc      231
                      Met Arg Ala Asn Asp Ala Leu Gln Val Leu Gly
                      1               5                   10

ttg ctt ttc agc ctg gcc cgg ggc tcc gag gtg ggc aac tct cag gca      279
Leu Leu Phe Ser Leu Ala Arg Gly Ser Glu Val Gly Asn Ser Gln Ala
            15                  20                  25

gtg tgt cct ggg act ctg aat ggc ctg agt gtg acc ggc gat gct gag      327
Val Cys Pro Gly Thr Leu Asn Gly Leu Ser Val Thr Gly Asp Ala Glu
        30                  35                  40

aac caa tac cag aca ctg tac aag ctc tac gag agg tgt gag gtg gtg      375
Asn Gln Tyr Gln Thr Leu Tyr Lys Leu Tyr Glu Arg Cys Glu Val Val
    45                  50                  55

atg ggg aac ctt gag att gtg ctc acg gga cac aat gcc gac ctc tcc      423
Met Gly Asn Leu Glu Ile Val Leu Thr Gly His Asn Ala Asp Leu Ser
60                  65                  70                  75

ttc ctg cag tgg att cga gaa gtg aca ggc tat gtc ctc gtg gcc atg      471
Phe Leu Gln Trp Ile Arg Glu Val Thr Gly Tyr Val Leu Val Ala Met
                80                  85                  90
```

```
aat gaa ttc tct act cta cca ttg ccc aac ctc cgc gtg gtg cga ggg      519
Asn Glu Phe Ser Thr Leu Pro Leu Pro Asn Leu Arg Val Val Arg Gly
            95                  100                 105

acc cag gtc tac gat ggg aag ttt gcc atc ttc gtc atg ttg aac tat      567
Thr Gln Val Tyr Asp Gly Lys Phe Ala Ile Phe Val Met Leu Asn Tyr
        110                 115                 120

aac acc aac tcc agc cac gct ctg cgc cag ctc cgc ttg act cag ctc      615
Asn Thr Asn Ser Ser His Ala Leu Arg Gln Leu Arg Leu Thr Gln Leu
    125                 130                 135

acc gag att ctg tca ggg ggt gtt tat att gag aag aac gat aag ctt      663
Thr Glu Ile Leu Ser Gly Gly Val Tyr Ile Glu Lys Asn Asp Lys Leu
140                 145                 150                 155

tgt cac atg gac aca att gac tgg agg gac atc gtg agg gac cga gat      711
Cys His Met Asp Thr Ile Asp Trp Arg Asp Ile Val Arg Asp Arg Asp
                160                 165                 170

gct gag ata gtg gtg aag gac aat ggc aga agc tgt ccc ccc tgt cat      759
Ala Glu Ile Val Val Lys Asp Asn Gly Arg Ser Cys Pro Pro Cys His
            175                 180                 185

gag gtt tgc aag ggg cga tgc tgg ggt cct gga tca gaa gac tgc cag      807
Glu Val Cys Lys Gly Arg Cys Trp Gly Pro Gly Ser Glu Asp Cys Gln
        190                 195                 200

aca ttg acc aag acc atc tgt gct cct cag tgt aat ggt cac tgc ttt      855
Thr Leu Thr Lys Thr Ile Cys Ala Pro Gln Cys Asn Gly His Cys Phe
    205                 210                 215

ggg ccc aac ccc aac cag tgc tgc cat gat gag tgt gcc ggg ggc tgc      903
Gly Pro Asn Pro Asn Gln Cys Cys His Asp Glu Cys Ala Gly Gly Cys
220                 225                 230                 235

tca ggc cct cag gac aca gac tgc ttt gcc tgc cgg cac ttc aat gac      951
Ser Gly Pro Gln Asp Thr Asp Cys Phe Ala Cys Arg His Phe Asn Asp
                240                 245                 250

agt gga gcc tgt gta cct cgc tgt cca cag cct ctt gtc tac aac aag      999
Ser Gly Ala Cys Val Pro Arg Cys Pro Gln Pro Leu Val Tyr Asn Lys
            255                 260                 265

cta act ttc cag ctg gaa ccc aat ccc cac acc aag tat cag tat gga     1047
Leu Thr Phe Gln Leu Glu Pro Asn Pro His Thr Lys Tyr Gln Tyr Gly
        270                 275                 280

gga gtt tgt gta gcc agc tgt ccc cat aac ttt gtg gtg gat caa aca     1095
Gly Val Cys Val Ala Ser Cys Pro His Asn Phe Val Val Asp Gln Thr
    285                 290                 295

tcc tgt gtc agg gcc tgt cct cct gac aag atg gaa gta gat aaa aat     1143
Ser Cys Val Arg Ala Cys Pro Pro Asp Lys Met Glu Val Asp Lys Asn
300                 305                 310                 315

ggg ctc aag atg tgt gag cct tgt ggg gga cta tgt ccc aaa gcc tgt     1191
Gly Leu Lys Met Cys Glu Pro Cys Gly Gly Leu Cys Pro Lys Ala Cys
                320                 325                 330

gag gga aca ggc tct ggg agc cgc ttc cag act gtg gac tcg agc aac     1239
Glu Gly Thr Gly Ser Gly Ser Arg Phe Gln Thr Val Asp Ser Ser Asn
            335                 340                 345
```

```
att gat gga ttt gtg aac tgc acc aag atc ctg ggc aac ctg gac ttt     1287
Ile Asp Gly Phe Val Asn Cys Thr Lys Ile Leu Gly Asn Leu Asp Phe
        350                 355                 360

ctg atc acc ggc ctc aat gga gac ccc tgg cac aag atc cct gcc ctg     1335
Leu Ile Thr Gly Leu Asn Gly Asp Pro Trp His Lys Ile Pro Ala Leu
    365                 370                 375

gac cca gag aag ctc aat gtc ttc cgg aca gta cgg gag atc aca ggt     1383
Asp Pro Glu Lys Leu Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly
380                 385                 390                 395

tac ctg aac atc cag tcc tgg ccg ccc cac atg cac aac ttc agt gtt     1431
Tyr Leu Asn Ile Gln Ser Trp Pro Pro His Met His Asn Phe Ser Val
                400                 405                 410

ttt tcc aat ttg aca acc att gga ggc aga agc ctc tac aac cgg ggc     1479
Phe Ser Asn Leu Thr Thr Ile Gly Gly Arg Ser Leu Tyr Asn Arg Gly
            415                 420                 425

ttc tca ttg ttg atc atg aag aac ttg aat gtc aca tct ctg ggc ttc     1527
Phe Ser Leu Leu Ile Met Lys Asn Leu Asn Val Thr Ser Leu Gly Phe
        430                 435                 440

cga tcc ctg aag gaa att agt gct ggg cgt atc tat ata agt gcc aat     1575
Arg Ser Leu Lys Glu Ile Ser Ala Gly Arg Ile Tyr Ile Ser Ala Asn
    445                 450                 455

agg cag ctc tgc tac cac cac tct ttg aac tgg acc aag gtg ctt cgg     1623
Arg Gln Leu Cys Tyr His His Ser Leu Asn Trp Thr Lys Val Leu Arg
460                 465                 470                 475

ggg cct acg gaa gag cga cta gac atc aag cat aat cgg ccg cgc aga     1671
Gly Pro Thr Glu Glu Arg Leu Asp Ile Lys His Asn Arg Pro Arg Arg
                480                 485                 490

gac tgc gtg gca gag ggc aaa gtg tgt gac cca ctg tgc tcc tct ggg     1719
Asp Cys Val Ala Glu Gly Lys Val Cys Asp Pro Leu Cys Ser Ser Gly
            495                 500                 505

gga tgc tgg ggc cca ggc cct ggt cag tgc ttg tcc tgt cga aat tat     1767
Gly Cys Trp Gly Pro Gly Pro Gly Gln Cys Leu Ser Cys Arg Asn Tyr
        510                 515                 520

agc cga gga ggt gtc tgt gtg acc cac tgc aac ttt ctg aat ggg gag     1815
Ser Arg Gly Gly Val Cys Val Thr His Cys Asn Phe Leu Asn Gly Glu
    525                 530                 535

cct cga gaa ttt gcc cat gag gcc gaa tgc ttc tcc tgc cac ccg gaa     1863
Pro Arg Glu Phe Ala His Glu Ala Glu Cys Phe Ser Cys His Pro Glu
540                 545                 550                 555

tgc caa ccc atg gag ggc act gcc aca tgc aat ggc tcg ggc tct gat     1911
Cys Gln Pro Met Glu Gly Thr Ala Thr Cys Asn Gly Ser Gly Ser Asp
                560                 565                 570

act tgt gct caa tgt gcc cat ttt cga gat ggg ccc cac tgt gtg agc     1959
Thr Cys Ala Gln Cys Ala His Phe Arg Asp Gly Pro His Cys Val Ser
            575                 580                 585

agc tgc ccc cat gga gtc cta ggt gcc aag ggc cca atc tac aag tac     2007
Ser Cys Pro His Gly Val Leu Gly Ala Lys Gly Pro Ile Tyr Lys Tyr
        590                 595                 600
```

```
cca gat gtt cag aat gaa tgt cgg ccc tgc cat gag aac tgc acc cag      2055
Pro Asp Val Gln Asn Glu Cys Arg Pro Cys His Glu Asn Cys Thr Gln
    605                 610                 615

ggg tgt aaa gga cca gag ctt caa gac tgt tta gga caa aca ctg gtg      2103
Gly Cys Lys Gly Pro Glu Leu Gln Asp Cys Leu Gly Gln Thr Leu Val
620                 625                 630                 635

ctg atc ggc aaa acc cat ctg aca atg gct ttg aca gtg ata gca gga      2151
Leu Ile Gly Lys Thr His Leu Thr Met Ala Leu Thr Val Ile Ala Gly
                640                 645                 650

ttg gta gtg att ttc atg atg ctg ggc ggc act ttt ctc tac tgg cgt      2199
Leu Val Val Ile Phe Met Met Leu Gly Gly Thr Phe Leu Tyr Trp Arg
            655                 660                 665

ggg cgc cgg att cag aat aaa agg gct atg agg cga tac ttg gaa cgg      2247
Gly Arg Arg Ile Gln Asn Lys Arg Ala Met Arg Arg Tyr Leu Glu Arg
        670                 675                 680

ggt gag agc ata gag cct ctg gac ccc agt gag aag gct aac aaa gtc      2295
Gly Glu Ser Ile Glu Pro Leu Asp Pro Ser Glu Lys Ala Asn Lys Val
    685                 690                 695

ttg gcc aga atc ttc aaa gag aca gag cta agg aag ctt aaa gtg ctt      2343
Leu Ala Arg Ile Phe Lys Glu Thr Glu Leu Arg Lys Leu Lys Val Leu
700                 705                 710                 715

ggc tcg ggt gtc ttt gga act gtg cac aaa gga gtg tgg atc cct gag      2391
Gly Ser Gly Val Phe Gly Thr Val His Lys Gly Val Trp Ile Pro Glu
                720                 725                 730

ggt gaa tca atc aag att cca gtc tgc att aaa gtc att gag gac aag      2439
Gly Glu Ser Ile Lys Ile Pro Val Cys Ile Lys Val Ile Glu Asp Lys
            735                 740                 745

agt gga cgg cag agt ttt caa gct gtg aca gat cat atg ctg gcc att      2487
Ser Gly Arg Gln Ser Phe Gln Ala Val Thr Asp His Met Leu Ala Ile
        750                 755                 760

ggc agc ctg gac cat gcc cac att gta agg ctg ctg gga cta tgc cca      2535
Gly Ser Leu Asp His Ala His Ile Val Arg Leu Leu Gly Leu Cys Pro
    765                 770                 775

ggg tca tct ctg cag ctt gtc act caa tat ttg cct ctg ggt tct ctg      2583
Gly Ser Ser Leu Gln Leu Val Thr Gln Tyr Leu Pro Leu Gly Ser Leu
780                 785                 790                 795

ctg gat cat gtg aga caa cac cgg ggg gca ctg ggg cca cag ctg ctg      2631
Leu Asp His Val Arg Gln His Arg Gly Ala Leu Gly Pro Gln Leu Leu
                800                 805                 810

ctc aac tgg gga gta caa att gcc aag gga atg tac tac ctt gag gaa      2679
Leu Asn Trp Gly Val Gln Ile Ala Lys Gly Met Tyr Tyr Leu Glu Glu
            815                 820                 825

cat ggt atg gtg cat aga aac ctg gct gcc cga aac gtg cta ctc aag      2727
His Gly Met Val His Arg Asn Leu Ala Ala Arg Asn Val Leu Leu Lys
        830                 835                 840

tca ccc agt cag gtt cag gtg gca gat ttt ggt gtg gct gac ctg ctg      2775
Ser Pro Ser Gln Val Gln Val Ala Asp Phe Gly Val Ala Asp Leu Leu
    845                 850                 855
```

```
cct cct gat gat aag cag ctg cta tac agt gag gcc aag act cca att     2823
Pro Pro Asp Asp Lys Gln Leu Leu Tyr Ser Glu Ala Lys Thr Pro Ile
860                 865                 870                 875

aag tgg atg gcc ctt gag agt atc cac ttt ggg aaa tac aca cac cag     2871
Lys Trp Met Ala Leu Glu Ser Ile His Phe Gly Lys Tyr Thr His Gln
                880                 885                 890

agt gat gtc tgg agc tat ggt gtg aca gtt tgg gag ttg atg acc ttc     2919
Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe
            895                 900                 905

ggg gca gag ccc tat gca ggg cta cga ttg gct gaa gta cca gac ctg     2967
Gly Ala Glu Pro Tyr Ala Gly Leu Arg Leu Ala Glu Val Pro Asp Leu
        910                 915                 920

cta gag aag ggg gag cgg ttg gca cag ccc cag atc tgc aca att gat     3015
Leu Glu Lys Gly Glu Arg Leu Ala Gln Pro Gln Ile Cys Thr Ile Asp
    925                 930                 935

gtc tac atg gtg atg gtc aag tgt tgg atg att gat gag aac att cgc     3063
Val Tyr Met Val Met Val Lys Cys Trp Met Ile Asp Glu Asn Ile Arg
940                 945                 950                 955

cca acc ttt aaa gaa cta gcc aat gag ttc acc agg atg gcc cga gac     3111
Pro Thr Phe Lys Glu Leu Ala Asn Glu Phe Thr Arg Met Ala Arg Asp
                960                 965                 970

cca cca cgg tat ctg gtc ata aag aga gag agt ggg cct gga ata gcc     3159
Pro Pro Arg Tyr Leu Val Ile Lys Arg Glu Ser Gly Pro Gly Ile Ala
            975                 980                 985

cct ggg cca gag ccc cat ggt ctg aca aac aag aag cta  gag gaa gta    3207
Pro Gly Pro Glu Pro His Gly Leu Thr Asn Lys Lys Leu  Glu Glu Val
        990                 995                 1000

gag ctg  gag cca gaa cta gac  cta gac cta gac ttg  gaa gca gag      3252
Glu Leu  Glu Pro Glu Leu Asp  Leu Asp Leu Asp Leu  Glu Ala Glu
    1005                1010                1015

gag gac  aac ctg gca acc acc  aca ctg ggc tcc gcc  ctc agc cta      3297
Glu Asp  Asn Leu Ala Thr Thr  Thr Leu Gly Ser Ala  Leu Ser Leu
    1020                1025                1030

cca gtt  gga aca ctt aat cgg  cca cgt ggg agc cag  agc ctt tta      3342
Pro Val  Gly Thr Leu Asn Arg  Pro Arg Gly Ser Gln  Ser Leu Leu
    1035                1040                1045

agt cca  tca tct gga tac atg  ccc atg aac cag ggt  aat ctt ggg      3387
Ser Pro  Ser Ser Gly Tyr Met  Pro Met Asn Gln Gly  Asn Leu Gly
    1050                1055                1060

ggg tct  tgc cag gag tct gca  gtt tct ggg agc agt  gaa cgg tgc      3432
Gly Ser  Cys Gln Glu Ser Ala  Val Ser Gly Ser Ser  Glu Arg Cys
    1065                1070                1075

ccc cgt  cca gtc tct cta cac  cca atg cca cgg gga  tgc ctg gca      3477
Pro Arg  Pro Val Ser Leu His  Pro Met Pro Arg Gly  Cys Leu Ala
    1080                1085                1090

tca gag  tca tca gag ggg cat  gta aca ggc tct gag  gct gag ctc      3522
Ser Glu  Ser Ser Glu Gly His  Val Thr Gly Ser Glu  Ala Glu Leu
    1095                1100                1105
```

```
cag gag  aaa gtg tca atg tgt  aga agc cgg agc agg  agc cgg agc      3567
Gln Glu  Lys Val Ser Met Cys  Arg Ser Arg Ser Arg  Ser Arg Ser
    1110                 1115                 1120

cca cgg  cca cgc gga gat agc  gcc tac cat tcc cag  cgc cac agt      3612
Pro Arg  Pro Arg Gly Asp Ser  Ala Tyr His Ser Gln  Arg His Ser
    1125                 1130                 1135

ctg ctg  act cct gtt acc cca  ctc tcc cca ccc ggg  tta gag gaa      3657
Leu Leu  Thr Pro Val Thr Pro  Leu Ser Pro Pro Gly  Leu Glu Glu
    1140                 1145                 1150

gag gat  gtc aac ggt tat gtc  atg cca gat aca cac  ctc aaa ggt      3702
Glu Asp  Val Asn Gly Tyr Val  Met Pro Asp Thr His  Leu Lys Gly
    1155                 1160                 1165

act ccc  tcc tcc cgg gaa ggc  acc ctt tct tca gtg  ggt ctc agt      3747
Thr Pro  Ser Ser Arg Glu Gly  Thr Leu Ser Ser Val  Gly Leu Ser
    1170                 1175                 1180

tct gtc  ctg ggt act gaa gaa  gaa gat gaa gat gag  gag tat gaa      3792
Ser Val  Leu Gly Thr Glu Glu  Glu Asp Glu Asp Glu  Glu Tyr Glu
    1185                 1190                 1195

tac atg  aac cgg agg aga agg  cac agt cca cct cat  ccc cct agg      3837
Tyr Met  Asn Arg Arg Arg Arg  His Ser Pro Pro His  Pro Pro Arg
    1200                 1205                 1210

cca agt  tcc ctt gag gag ctg  ggt tat gag tac atg  gat gtg ggg      3882
Pro Ser  Ser Leu Glu Glu Leu  Gly Tyr Glu Tyr Met  Asp Val Gly
    1215                 1220                 1225

tca gac  ctc agt gcc tct ctg  ggc agc aca cag agt  tgc cca ctc      3927
Ser Asp  Leu Ser Ala Ser Leu  Gly Ser Thr Gln Ser  Cys Pro Leu
    1230                 1235                 1240

cac cct  gta ccc atc atg ccc  act gca ggc aca act  cca gat gaa      3972
His Pro  Val Pro Ile Met Pro  Thr Ala Gly Thr Thr  Pro Asp Glu
    1245                 1250                 1255

gac tat  gaa tat atg aat cgg  caa cga gat gga ggt  ggt cct ggg      4017
Asp Tyr  Glu Tyr Met Asn Arg  Gln Arg Asp Gly Gly  Gly Pro Gly
    1260                 1265                 1270

ggt gat  tat gca gcc atg ggg  gcc tgc cca gca tct  gag caa ggg      4062
Gly Asp  Tyr Ala Ala Met Gly  Ala Cys Pro Ala Ser  Glu Gln Gly
    1275                 1280                 1285

tat gaa  gag atg aga gct ttt  cag ggg cct gga cat  cag gcc ccc      4107
Tyr Glu  Glu Met Arg Ala Phe  Gln Gly Pro Gly His  Gln Ala Pro
    1290                 1295                 1300

cat gtc  cat tat gcc cgc cta  aaa act cta cgt agc  tta gag gct      4152
His Val  His Tyr Ala Arg Leu  Lys Thr Leu Arg Ser  Leu Glu Ala
    1305                 1310                 1315

aca gac  tct gcc ttt gat aac  cct gat tac tgg cat  agc agg ctt      4197
Thr Asp  Ser Ala Phe Asp Asn  Pro Asp Tyr Trp His  Ser Arg Leu
    1320                 1325                 1330

ttc ccc  aag gct aat gcc cag  aga acg taa ctcctgctcc ctgtggcact     4247
Phe Pro  Lys Ala Asn Ala Gln  Arg Thr
    1335                 1340
```

cagggagcat ttaatggcag ctagtgcctt tagagggtac cgtcttctcc ctattccctc 4307

tctctcccag gtcccagccc cttttcccca gtcccagaca attccattca atctttggag 4367

gcttttaaac attttgacac aaaattctta tggtatgtag ccagctgtgc actttcttct 4427

ctttcccaac cccaggaaag gttttcctta ttttgtgtgc tttcccagtc ccattcctca 4487

gcttcttcac aggcactcct ggagatatga aggattactc tccatatccc ttcctctcag 4547

gctcttgact acttggaact aggctcttat gtgtgccttt gtttcccatc agactgtcaa 4607

gaagaggaaa gggaggaaac ctagcagagg aaagtgtaat tttggtttat gactcttaac 4667

cccctagaaa gacagaagct taaaatctgt gaagaaagag gttaggagta gatattgatt 4727

actatcataa ttcagcactt aactatgagc caggcatcat actaaacttc acctacatta 4787

tctcacttag tcctttatca tccttaaaac aattctgtga catacatatt atctcatttt 4847

acacaaaggg aagtcgggca tggtggctca tgcctgtaat ctcagcactt tgggaggctg 4907

aggcagaagg attacctgag gcaaggagtt tgagaccagc ttagccaaca tagtaagacc 4967

cccatctc 4975

<210> 14
<211> 1342
<212> PRT
<213> Homo sapiens

<400> 14

Met Arg Ala Asn Asp Ala Leu Gln Val Leu Gly Leu Leu Phe Ser Leu
1 5 10 15

Ala Arg Gly Ser Glu Val Gly Asn Ser Gln Ala Val Cys Pro Gly Thr
20 25 30

Leu Asn Gly Leu Ser Val Thr Gly Asp Ala Glu Asn Gln Tyr Gln Thr
35 40 45

Leu Tyr Lys Leu Tyr Glu Arg Cys Glu Val Val Met Gly Asn Leu Glu
50 55 60

Ile Val Leu Thr Gly His Asn Ala Asp Leu Ser Phe Leu Gln Trp Ile
65 70 75 80

Arg Glu Val Thr Gly Tyr Val Leu Val Ala Met Asn Glu Phe Ser Thr
85 90 95

Leu Pro Leu Pro Asn Leu Arg Val Val Arg Gly Thr Gln Val Tyr Asp
100 105 110

Gly Lys Phe Ala Ile Phe Val Met Leu Asn Tyr Asn Thr Asn Ser Ser

115 120 125

His Ala Leu Arg Gln Leu Arg Leu Thr Gln Leu Thr Glu Ile Leu Ser
130 135 140

Gly Gly Val Tyr Ile Glu Lys Asn Asp Lys Leu Cys His Met Asp Thr
145 150 155 160

Ile Asp Trp Arg Asp Ile Val Arg Asp Arg Asp Ala Glu Ile Val Val
165 170 175

Lys Asp Asn Gly Arg Ser Cys Pro Pro Cys His Glu Val Cys Lys Gly
180 185 190

Arg Cys Trp Gly Pro Gly Ser Glu Asp Cys Gln Thr Leu Thr Lys Thr
195 200 205

Ile Cys Ala Pro Gln Cys Asn Gly His Cys Phe Gly Pro Asn Pro Asn
210 215 220

Gln Cys Cys His Asp Glu Cys Ala Gly Gly Cys Ser Gly Pro Gln Asp
225 230 235 240

Thr Asp Cys Phe Ala Cys Arg His Phe Asn Asp Ser Gly Ala Cys Val
245 250 255

Pro Arg Cys Pro Gln Pro Leu Val Tyr Asn Lys Leu Thr Phe Gln Leu
260 265 270

Glu Pro Asn Pro His Thr Lys Tyr Gln Tyr Gly Gly Val Cys Val Ala
275 280 285

Ser Cys Pro His Asn Phe Val Val Asp Gln Thr Ser Cys Val Arg Ala
290 295 300

Cys Pro Pro Asp Lys Met Glu Val Asp Lys Asn Gly Leu Lys Met Cys
305 310 315 320

Glu Pro Cys Gly Gly Leu Cys Pro Lys Ala Cys Glu Gly Thr Gly Ser
325 330 335

Gly Ser Arg Phe Gln Thr Val Asp Ser Ser Asn Ile Asp Gly Phe Val
340 345 350

Asn Cys Thr Lys Ile Leu Gly Asn Leu Asp Phe Leu Ile Thr Gly Leu
355 360 365

Asn Gly Asp Pro Trp His Lys Ile Pro Ala Leu Asp Pro Glu Lys Leu

370 375 380

Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly Tyr Leu Asn Ile Gln
385 390 395 400

Ser Trp Pro Pro His Met His Asn Phe Ser Val Phe Ser Asn Leu Thr
405 410 415

Thr Ile Gly Gly Arg Ser Leu Tyr Asn Arg Gly Phe Ser Leu Leu Ile
420 425 430

Met Lys Asn Leu Asn Val Thr Ser Leu Gly Phe Arg Ser Leu Lys Glu
435 440 445

Ile Ser Ala Gly Arg Ile Tyr Ile Ser Ala Asn Arg Gln Leu Cys Tyr
450 455 460

His His Ser Leu Asn Trp Thr Lys Val Leu Arg Gly Pro Thr Glu Glu
465 470 475 480

Arg Leu Asp Ile Lys His Asn Arg Pro Arg Arg Asp Cys Val Ala Glu
485 490 495

Gly Lys Val Cys Asp Pro Leu Cys Ser Ser Gly Gly Cys Trp Gly Pro
500 505 510

Gly Pro Gly Gln Cys Leu Ser Cys Arg Asn Tyr Ser Arg Gly Gly Val
515 520 525

Cys Val Thr His Cys Asn Phe Leu Asn Gly Glu Pro Arg Glu Phe Ala
530 535 540

His Glu Ala Glu Cys Phe Ser Cys His Pro Glu Cys Gln Pro Met Glu
545 550 555 560

Gly Thr Ala Thr Cys Asn Gly Ser Gly Ser Asp Thr Cys Ala Gln Cys
565 570 575

Ala His Phe Arg Asp Gly Pro His Cys Val Ser Ser Cys Pro His Gly
580 585 590

Val Leu Gly Ala Lys Gly Pro Ile Tyr Lys Tyr Pro Asp Val Gln Asn
595 600 605

Glu Cys Arg Pro Cys His Glu Asn Cys Thr Gln Gly Cys Lys Gly Pro
610 615 620

Glu Leu Gln Asp Cys Leu Gly Gln Thr Leu Val Leu Ile Gly Lys Thr

625 630 635 640

His Leu Thr Met Ala Leu Thr Val Ile Ala Gly Leu Val Val Ile Phe
645 650 655

Met Met Leu Gly Gly Thr Phe Leu Tyr Trp Arg Gly Arg Arg Ile Gln
660 665 670

Asn Lys Arg Ala Met Arg Arg Tyr Leu Glu Arg Gly Glu Ser Ile Glu
675 680 685

Pro Leu Asp Pro Ser Glu Lys Ala Asn Lys Val Leu Ala Arg Ile Phe
690 695 700

Lys Glu Thr Glu Leu Arg Lys Leu Lys Val Leu Gly Ser Gly Val Phe
705 710 715 720

Gly Thr Val His Lys Gly Val Trp Ile Pro Glu Gly Glu Ser Ile Lys
725 730 735

Ile Pro Val Cys Ile Lys Val Ile Glu Asp Lys Ser Gly Arg Gln Ser
740 745 750

Phe Gln Ala Val Thr Asp His Met Leu Ala Ile Gly Ser Leu Asp His
755 760 765

Ala His Ile Val Arg Leu Leu Gly Leu Cys Pro Gly Ser Ser Leu Gln
770 775 780

Leu Val Thr Gln Tyr Leu Pro Leu Gly Ser Leu Leu Asp His Val Arg
785 790 795 800

Gln His Arg Gly Ala Leu Gly Pro Gln Leu Leu Leu Asn Trp Gly Val
805 810 815

Gln Ile Ala Lys Gly Met Tyr Tyr Leu Glu Glu His Gly Met Val His
820 825 830

Arg Asn Leu Ala Ala Arg Asn Val Leu Leu Lys Ser Pro Ser Gln Val
835 840 845

Gln Val Ala Asp Phe Gly Val Ala Asp Leu Leu Pro Pro Asp Asp Lys
850 855 860

Gln Leu Leu Tyr Ser Glu Ala Lys Thr Pro Ile Lys Trp Met Ala Leu
865 870 875 880

Glu Ser Ile His Phe Gly Lys Tyr Thr His Gln Ser Asp Val Trp Ser

885 890 895

Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe Gly Ala Glu Pro Tyr
900 905 910

Ala Gly Leu Arg Leu Ala Glu Val Pro Asp Leu Leu Glu Lys Gly Glu
915 920 925

Arg Leu Ala Gln Pro Gln Ile Cys Thr Ile Asp Val Tyr Met Val Met
930 935 940

Val Lys Cys Trp Met Ile Asp Glu Asn Ile Arg Pro Thr Phe Lys Glu
945 950 955 960

Leu Ala Asn Glu Phe Thr Arg Met Ala Arg Asp Pro Pro Arg Tyr Leu
965 970 975

Val Ile Lys Arg Glu Ser Gly Pro Gly Ile Ala Pro Gly Pro Glu Pro
980 985 990

His Gly Leu Thr Asn Lys Lys Leu Glu Glu Val Glu Leu Glu Pro Glu
995 1000 1005

Leu Asp Leu Asp Leu Asp Leu Glu Ala Glu Glu Asp Asn Leu Ala
1010 1015 1020

Thr Thr Thr Leu Gly Ser Ala Leu Ser Leu Pro Val Gly Thr Leu
1025 1030 1035

Asn Arg Pro Arg Gly Ser Gln Ser Leu Leu Ser Pro Ser Ser Gly
1040 1045 1050

Tyr Met Pro Met Asn Gln Gly Asn Leu Gly Gly Ser Cys Gln Glu
1055 1060 1065

Ser Ala Val Ser Gly Ser Ser Glu Arg Cys Pro Arg Pro Val Ser
1070 1075 1080

Leu His Pro Met Pro Arg Gly Cys Leu Ala Ser Glu Ser Ser Glu
1085 1090 1095

Gly His Val Thr Gly Ser Glu Ala Glu Leu Gln Glu Lys Val Ser
1100 1105 1110

Met Cys Arg Ser Arg Ser Arg Ser Arg Ser Pro Arg Pro Arg Gly
1115 1120 1125

Asp Ser Ala Tyr His Ser Gln Arg His Ser Leu Leu Thr Pro Val

1130 1135 1140

Thr Pro Leu Ser Pro Pro Gly Leu Glu Glu Glu Asp Val Asn Gly
1145 1150 1155

Tyr Val Met Pro Asp Thr His Leu Lys Gly Thr Pro Ser Ser Arg
1160 1165 1170

Glu Gly Thr Leu Ser Ser Val Gly Leu Ser Ser Val Leu Gly Thr
1175 1180 1185

Glu Glu Glu Asp Glu Asp Glu Glu Tyr Glu Tyr Met Asn Arg Arg
1190 1195 1200

Arg Arg His Ser Pro Pro His Pro Pro Arg Pro Ser Ser Leu Glu
1205 1210 1215

Glu Leu Gly Tyr Glu Tyr Met Asp Val Gly Ser Asp Leu Ser Ala
1220 1225 1230

Ser Leu Gly Ser Thr Gln Ser Cys Pro Leu His Pro Val Pro Ile
1235 1240 1245

Met Pro Thr Ala Gly Thr Thr Pro Asp Glu Asp Tyr Glu Tyr Met
1250 1255 1260

Asn Arg Gln Arg Asp Gly Gly Gly Pro Gly Gly Asp Tyr Ala Ala
1265 1270 1275

Met Gly Ala Cys Pro Ala Ser Glu Gln Gly Tyr Glu Glu Met Arg
1280 1285 1290

Ala Phe Gln Gly Pro Gly His Gln Ala Pro His Val His Tyr Ala
1295 1300 1305

Arg Leu Lys Thr Leu Arg Ser Leu Glu Ala Thr Asp Ser Ala Phe
1310 1315 1320

Asp Asn Pro Asp Tyr Trp His Ser Arg Leu Phe Pro Lys Ala Asn
1325 1330 1335

Ala Gln Arg Thr
1340

<210> 15
<211> 5484
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (34)..(3960)
<223> HER-4 coding sequence

<400> 15

```
aattgtcagc acgggatctg agacttccaa aaa atg aag ccg gcg aca gga ctt      54
                                      Met Lys Pro Ala Thr Gly Leu
                                      1               5

tgg gtc tgg gtg agc ctt ctc gtg gcg gcg ggg acc gtc cag ccc agc     102
Trp Val Trp Val Ser Leu Leu Val Ala Ala Gly Thr Val Gln Pro Ser
        10                  15                  20

gat tct cag tca gtg tgt gca gga acg gag aat aaa ctg agc tct ctc     150
Asp Ser Gln Ser Val Cys Ala Gly Thr Glu Asn Lys Leu Ser Ser Leu
    25                  30                  35

tct gac ctg gaa cag cag tac cga gcc ttg cgc aag tac tat gaa aac     198
Ser Asp Leu Glu Gln Gln Tyr Arg Ala Leu Arg Lys Tyr Tyr Glu Asn
40                  45                  50                  55

tgt gag gtt gtc atg ggc aac ctg gag ata acc agc att gag cac aac     246
Cys Glu Val Val Met Gly Asn Leu Glu Ile Thr Ser Ile Glu His Asn
                60                  65                  70

cgg gac ctc tcc ttc ctg cgg tct gtt cga gaa gtc aca ggc tac gtg     294
Arg Asp Leu Ser Phe Leu Arg Ser Val Arg Glu Val Thr Gly Tyr Val
            75                  80                  85

tta gtg gct ctt aat cag ttt cgt tac ctg cct ctg gag aat tta cgc     342
Leu Val Ala Leu Asn Gln Phe Arg Tyr Leu Pro Leu Glu Asn Leu Arg
        90                  95                  100

att att cgt ggg aca aaa ctt tat gag gat cga tat gcc ttg gca ata     390
Ile Ile Arg Gly Thr Lys Leu Tyr Glu Asp Arg Tyr Ala Leu Ala Ile
    105                 110                 115

ttt tta aac tac aga aaa gat gga aac ttt gga ctt caa gaa ctt gga     438
Phe Leu Asn Tyr Arg Lys Asp Gly Asn Phe Gly Leu Gln Glu Leu Gly
120                 125                 130                 135

tta aag aac ttg aca gaa atc cta aat ggt gga gtc tat gta gac cag     486
Leu Lys Asn Leu Thr Glu Ile Leu Asn Gly Gly Val Tyr Val Asp Gln
                140                 145                 150

aac aaa ttc ctt tgt tat gca gac acc att cat tgg caa gat att gtt     534
Asn Lys Phe Leu Cys Tyr Ala Asp Thr Ile His Trp Gln Asp Ile Val
            155                 160                 165

cgg aac cca tgg cct tcc aac ttg act ctt gtg tca aca aat ggt agt     582
Arg Asn Pro Trp Pro Ser Asn Leu Thr Leu Val Ser Thr Asn Gly Ser
        170                 175                 180

tca gga tgt gga cgt tgc cat aag tcc tgt act ggc cgt tgc tgg gga     630
Ser Gly Cys Gly Arg Cys His Lys Ser Cys Thr Gly Arg Cys Trp Gly
    185                 190                 195

ccc aca gaa aat cat tgc cag act ttg aca agg acg gtg tgt gca gaa     678
Pro Thr Glu Asn His Cys Gln Thr Leu Thr Arg Thr Val Cys Ala Glu
200                 205                 210                 215

caa tgt gac ggc aga tgc tac gga cct tac gtc agt gac tgc tgc cat     726
```

```
Gln Cys Asp Gly Arg Cys Tyr Gly Pro Tyr Val Ser Asp Cys Cys His
                220                 225                 230

cga gaa tgt gct gga ggc tgc tca gga cct aag gac aca gac tgc ttt   774
Arg Glu Cys Ala Gly Gly Cys Ser Gly Pro Lys Asp Thr Asp Cys Phe
            235                 240                 245

gcc tgc atg aat ttc aat gac agt gga gca tgt gtt act cag tgt ccc   822
Ala Cys Met Asn Phe Asn Asp Ser Gly Ala Cys Val Thr Gln Cys Pro
        250                 255                 260

caa acc ttt gtc tac aat cca acc acc ttt caa ctg gag cac aat ttc   870
Gln Thr Phe Val Tyr Asn Pro Thr Thr Phe Gln Leu Glu His Asn Phe
    265                 270                 275

aat gca aag tac aca tat gga gca ttc tgt gtc aag aaa tgt cca cat   918
Asn Ala Lys Tyr Thr Tyr Gly Ala Phe Cys Val Lys Lys Cys Pro His
280                 285                 290                 295

aac ttt gtg gta gat tcc agt tct tgt gtg cgt gcc tgc cct agt tcc   966
Asn Phe Val Val Asp Ser Ser Ser Cys Val Arg Ala Cys Pro Ser Ser
                300                 305                 310

aag atg gaa gta gaa gaa aat ggg att aaa atg tgt aaa cct tgc act  1014
Lys Met Glu Val Glu Glu Asn Gly Ile Lys Met Cys Lys Pro Cys Thr
            315                 320                 325

gac att tgc cca aaa gct tgt gat ggc att ggc aca gga tca ttg atg  1062
Asp Ile Cys Pro Lys Ala Cys Asp Gly Ile Gly Thr Gly Ser Leu Met
        330                 335                 340

tca gct cag act gtg gat tcc agt aac att gac aaa ttc ata aac tgt  1110
Ser Ala Gln Thr Val Asp Ser Ser Asn Ile Asp Lys Phe Ile Asn Cys
    345                 350                 355

acc aag atc aat ggg aat ttg atc ttt cta gtc act ggt att cat ggg  1158
Thr Lys Ile Asn Gly Asn Leu Ile Phe Leu Val Thr Gly Ile His Gly
360                 365                 370                 375

gac cct tac aat gca att gaa gcc ata gac cca gag aaa ctg aac gtc  1206
Asp Pro Tyr Asn Ala Ile Glu Ala Ile Asp Pro Glu Lys Leu Asn Val
                380                 385                 390

ttt cgg aca gtc aga gag ata aca ggt ttc ctg aac ata cag tca tgg  1254
Phe Arg Thr Val Arg Glu Ile Thr Gly Phe Leu Asn Ile Gln Ser Trp
            395                 400                 405

cca cca aac atg act gac ttc agt gtt ttt tct aac ctg gtg acc att  1302
Pro Pro Asn Met Thr Asp Phe Ser Val Phe Ser Asn Leu Val Thr Ile
        410                 415                 420

ggt gga aga gta ctc tat agt ggc ctg tcc ttg ctt atc ctc aag caa  1350
Gly Gly Arg Val Leu Tyr Ser Gly Leu Ser Leu Leu Ile Leu Lys Gln
    425                 430                 435

cag ggc atc acc tct cta cag ttc cag tcc ctg aag gaa atc agc gca  1398
Gln Gly Ile Thr Ser Leu Gln Phe Gln Ser Leu Lys Glu Ile Ser Ala
440                 445                 450                 455

gga aac atc tat att act gac aac agc aac ctg tgt tat tat cat acc  1446
Gly Asn Ile Tyr Ile Thr Asp Asn Ser Asn Leu Cys Tyr Tyr His Thr
                460                 465                 470

att aac tgg aca aca ctc ttc agc aca atc aac cag aga ata gta atc  1494
```

```
Ile Asn Trp Thr Thr Leu Phe Ser Thr Ile Asn Gln Arg Ile Val Ile
            475                 480                 485

cgg gac aac aga aaa gct gaa aat tgt act gct gaa gga atg gtg tgc      1542
Arg Asp Asn Arg Lys Ala Glu Asn Cys Thr Ala Glu Gly Met Val Cys
        490                 495                 500

aac cat ctg tgt tcc agt gat ggc tgt tgg gga cct ggg cca gac caa      1590
Asn His Leu Cys Ser Ser Asp Gly Cys Trp Gly Pro Gly Pro Asp Gln
    505                 510                 515

tgt ctg tcg tgt cgc cgc ttc agt aga gga agg atc tgc ata gag tct      1638
Cys Leu Ser Cys Arg Arg Phe Ser Arg Gly Arg Ile Cys Ile Glu Ser
520                 525                 530                 535

tgt aac ctc tat gat ggt gaa ttt cgg gag ttt gag aat ggc tcc atc      1686
Cys Asn Leu Tyr Asp Gly Glu Phe Arg Glu Phe Glu Asn Gly Ser Ile
                540                 545                 550

tgt gtg gag tgt gac ccc cag tgt gag aag atg gaa gat ggc ctc ctc      1734
Cys Val Glu Cys Asp Pro Gln Cys Glu Lys Met Glu Asp Gly Leu Leu
            555                 560                 565

aca tgc cat gga ccg ggt cct gac aac tgt aca aag tgc tct cat ttt      1782
Thr Cys His Gly Pro Gly Pro Asp Asn Cys Thr Lys Cys Ser His Phe
        570                 575                 580

aaa gat ggc cca aac tgt gtg gaa aaa tgt cca gat ggc tta cag ggg      1830
Lys Asp Gly Pro Asn Cys Val Glu Lys Cys Pro Asp Gly Leu Gln Gly
    585                 590                 595

gca aac agt ttc att ttc aag tat gct gat cca gat cgg gag tgc cac      1878
Ala Asn Ser Phe Ile Phe Lys Tyr Ala Asp Pro Asp Arg Glu Cys His
600                 605                 610                 615

cca tgc cat cca aac tgc acc caa ggg tgt aac ggt ccc act agt cat      1926
Pro Cys His Pro Asn Cys Thr Gln Gly Cys Asn Gly Pro Thr Ser His
                620                 625                 630

gac tgc att tac tac cca tgg acg ggc cat tcc act tta cca caa cat      1974
Asp Cys Ile Tyr Tyr Pro Trp Thr Gly His Ser Thr Leu Pro Gln His
            635                 640                 645

gct aga act ccc ctg att gca gct gga gta att ggt ggg ctc ttc att      2022
Ala Arg Thr Pro Leu Ile Ala Ala Gly Val Ile Gly Gly Leu Phe Ile
        650                 655                 660

ctg gtc att gtg ggt ctg aca ttt gct gtt tat gtt aga agg aag agc      2070
Leu Val Ile Val Gly Leu Thr Phe Ala Val Tyr Val Arg Arg Lys Ser
    665                 670                 675

atc aaa aag aaa aga gcc ttg aga aga ttc ttg gaa aca gag ttg gtg      2118
Ile Lys Lys Lys Arg Ala Leu Arg Arg Phe Leu Glu Thr Glu Leu Val
680                 685                 690                 695

gaa cca tta act ccc agt ggc aca gca ccc aat caa gct caa ctt cgt      2166
Glu Pro Leu Thr Pro Ser Gly Thr Ala Pro Asn Gln Ala Gln Leu Arg
                700                 705                 710

att ttg aaa gaa act gag ctg aag agg gta aaa gtc ctt ggc tca ggt      2214
Ile Leu Lys Glu Thr Glu Leu Lys Arg Val Lys Val Leu Gly Ser Gly
            715                 720                 725

gct ttt gga acg gtt tat aaa ggt att tgg gta cct gaa gga gaa act      2262
```

```
Ala Phe Gly Thr Val Tyr Lys Gly Ile Trp Val Pro Glu Gly Glu Thr
        730                 735                 740

gtg aag att cct gtg gct att aag att ctt aat gag aca act ggt ccc     2310
Val Lys Ile Pro Val Ala Ile Lys Ile Leu Asn Glu Thr Thr Gly Pro
    745                 750                 755

aag gca aat gtg gag ttc atg gat gaa gct ctg atc atg gca agt atg     2358
Lys Ala Asn Val Glu Phe Met Asp Glu Ala Leu Ile Met Ala Ser Met
760                 765                 770                 775

gat cat cca cac cta gtc cgg ttg ctg ggt gtg tgt ctg agc cca acc     2406
Asp His Pro His Leu Val Arg Leu Leu Gly Val Cys Leu Ser Pro Thr
                780                 785                 790

atc cag ctg gtt act caa ctt atg ccc cat ggc tgc ctg ttg gag tat     2454
Ile Gln Leu Val Thr Gln Leu Met Pro His Gly Cys Leu Leu Glu Tyr
            795                 800                 805

gtc cac gag cac aag gat aac att gga tca caa ctg ctg ctt aac tgg     2502
Val His Glu His Lys Asp Asn Ile Gly Ser Gln Leu Leu Leu Asn Trp
        810                 815                 820

tgt gtc cag ata gct aag gga atg atg tac ctg gaa gaa aga cga ctc     2550
Cys Val Gln Ile Ala Lys Gly Met Met Tyr Leu Glu Glu Arg Arg Leu
    825                 830                 835

gtt cat cgg gat ttg gca gcc cgt aat gtc tta gtg aaa tct cca aac     2598
Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Lys Ser Pro Asn
840                 845                 850                 855

cat gtg aaa atc aca gat ttt ggg cta gcc aga ctc ttg gaa gga gat     2646
His Val Lys Ile Thr Asp Phe Gly Leu Ala Arg Leu Leu Glu Gly Asp
                860                 865                 870

gaa aaa gag tac aat gct gat gga gga aag atg cca att aaa tgg atg     2694
Glu Lys Glu Tyr Asn Ala Asp Gly Gly Lys Met Pro Ile Lys Trp Met
            875                 880                 885

gct ctg gag tgt ata cat tac agg aaa ttc acc cat cag agt gac gtt     2742
Ala Leu Glu Cys Ile His Tyr Arg Lys Phe Thr His Gln Ser Asp Val
        890                 895                 900

tgg agc tat gga gtt act ata tgg gaa ctg atg acc ttt gga gga aaa     2790
Trp Ser Tyr Gly Val Thr Ile Trp Glu Leu Met Thr Phe Gly Gly Lys
    905                 910                 915

ccc tat gat gga att cca acg cga gaa atc cct gat tta tta gag aaa     2838
Pro Tyr Asp Gly Ile Pro Thr Arg Glu Ile Pro Asp Leu Leu Glu Lys
920                 925                 930                 935

gga gaa cgt ttg cct cag cct ccc atc tgc act att gac gtt tac atg     2886
Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met
                940                 945                 950

gtc atg gtc aaa tgt tgg atg att gat gct gac agt aga cct aaa ttt     2934
Val Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro Lys Phe
            955                 960                 965

aag gaa ctg gct gct gag ttt tca agg atg gct cga gac cct caa aga     2982
Lys Glu Leu Ala Ala Glu Phe Ser Arg Met Ala Arg Asp Pro Gln Arg
        970                 975                 980

tac cta gtt att cag ggt gat gat cgt atg aag ctt ccc agt cca aat     3030
```

```
Tyr Leu Val Ile Gln Gly Asp Asp Arg Met Lys Leu Pro Ser Pro Asn
    985                 990                 995

gac  agc aag ttc ttt cag  aat ctc ttg gat gaa  gag gat ttg gaa      3075
Asp  Ser Lys Phe Phe Gln  Asn Leu Leu Asp Glu  Glu Asp Leu Glu
1000                1005                1010

gat  atg atg gat gct gag  gag tac ttg gtc cct  cag gct ttc aac      3120
Asp  Met Met Asp Ala Glu  Glu Tyr Leu Val Pro  Gln Ala Phe Asn
1015                1020                1025

atc  cca cct ccc atc tat  act tcc aga gca aga  att gac tcg aat      3165
Ile  Pro Pro Pro Ile Tyr  Thr Ser Arg Ala Arg  Ile Asp Ser Asn
1030                1035                1040

agg  agt gaa att gga cac  agc cct cct cct gcc  tac acc ccc atg      3210
Arg  Ser Glu Ile Gly His  Ser Pro Pro Pro Ala  Tyr Thr Pro Met
1045                1050                1055

tca  gga aac cag ttt gta  tac cga gat gga ggt  ttt gct gct gaa      3255
Ser  Gly Asn Gln Phe Val  Tyr Arg Asp Gly Gly  Phe Ala Ala Glu
1060                1065                1070

caa  gga gtg tct gtg ccc  tac aga gcc cca act  agc aca att cca      3300
Gln  Gly Val Ser Val Pro  Tyr Arg Ala Pro Thr  Ser Thr Ile Pro
1075                1080                1085

gaa  gct cct gtg gca cag  ggt gct act gct gag  att ttt gat gac      3345
Glu  Ala Pro Val Ala Gln  Gly Ala Thr Ala Glu  Ile Phe Asp Asp
1090                1095                1100

tcc  tgc tgt aat ggc acc  cta cgc aag cca gtg  gca ccc cat gtc      3390
Ser  Cys Cys Asn Gly Thr  Leu Arg Lys Pro Val  Ala Pro His Val
1105                1110                1115

caa  gag gac agt agc acc  cag agg tac agt gct  gac ccc acc gtg      3435
Gln  Glu Asp Ser Ser Thr  Gln Arg Tyr Ser Ala  Asp Pro Thr Val
1120                1125                1130

ttt  gcc cca gaa cgg agc  cca cga gga gag ctg  gat gag gaa ggt      3480
Phe  Ala Pro Glu Arg Ser  Pro Arg Gly Glu Leu  Asp Glu Glu Gly
1135                1140                1145

tac  atg act cct atg cga  gac aaa ccc aaa caa  gaa tac ctg aat      3525
Tyr  Met Thr Pro Met Arg  Asp Lys Pro Lys Gln  Glu Tyr Leu Asn
1150                1155                1160

cca  gtg gag gag aac cct  ttt gtt tct cgg aga  aaa aat gga gac      3570
Pro  Val Glu Glu Asn Pro  Phe Val Ser Arg Arg  Lys Asn Gly Asp
1165                1170                1175

ctt  caa gca ttg gat aat  ccc gaa tat cac aat  gca tcc aat ggt      3615
Leu  Gln Ala Leu Asp Asn  Pro Glu Tyr His Asn  Ala Ser Asn Gly
1180                1185                1190

cca  ccc aag gcc gag gat  gag tat gtg aat gag  cca ctg tac ctc      3660
Pro  Pro Lys Ala Glu Asp  Glu Tyr Val Asn Glu  Pro Leu Tyr Leu
1195                1200                1205

aac  acc ttt gcc aac acc  ttg gga aaa gct gag  tac ctg aag aac      3705
Asn  Thr Phe Ala Asn Thr  Leu Gly Lys Ala Glu  Tyr Leu Lys Asn
1210                1215                1220

aac  ata ctg tca atg cca  gag aag gcc aag aaa  gcg ttt gac aac      3750
```

```
Asn  Ile Leu Ser Met Pro  Glu Lys Ala Lys Lys  Ala Phe Asp Asn
1225                 1230                 1235

cct  gac tac tgg aac cac  agc ctg cca cct cgg  agc acc ctt cag      3795
Pro  Asp Tyr Trp Asn His  Ser Leu Pro Pro Arg  Ser Thr Leu Gln
1240                 1245                 1250

cac  cca gac tac ctg cag  gag tac agc aca aaa  tat ttt tat aaa      3840
His  Pro Asp Tyr Leu Gln  Glu Tyr Ser Thr Lys  Tyr Phe Tyr Lys
1255                 1260                 1265

cag  aat ggg cgg atc cgg  cct att gtg gca gag  aat cct gaa tac      3885
Gln  Asn Gly Arg Ile Arg  Pro Ile Val Ala Glu  Asn Pro Glu Tyr
1270                 1275                 1280

ctc  tct gag ttc tcc ctg  aag cca ggc act gtg  ctg ccg cct cca      3930
Leu  Ser Glu Phe Ser Leu  Lys Pro Gly Thr Val  Leu Pro Pro Pro
1285                 1290                 1295

cct  tac aga cac cgg aat  act gtg gtg taa gctcagttgt ggtttttag      3980
Pro  Tyr Arg His Arg Asn  Thr Val Val
1300                 1305

gtggagagac acacctgctc caatttcccc acccccctct ctttctctgg tggtcttcct   4040

tctaccccaa ggccagtagt tttgacactt cccagtggaa gatacagaga tgcaatgata   4100

gttatgtgct tacctaactt gaacattaga gggaaagact gaaagagaaa gataggagga   4160

accacaatgt ttcttcattt ctctgcatgg gttggtcagg agaatgaaac agctagagaa   4220

ggaccagaaa atgtaaggca atgctgccta ctatcaaact agctgtcact ttttttcttt   4280

ttctttttct ttctttgttt ctttcttcct cttctttttt ttttttttt taaagcagat   4340

ggttgaaaca cccatgctat ctgttcctat ctgcaggaac tgatgtgtgc atatttagca   4400

tccctggaaa tcataataaa gtttccatta gaacaaaaga ataacatttt ctataacata   4460

tgatagtgtc tgaaattgag aatccagttt ctttccccag cagtttctgt cctagcaagt   4520

aagaatggcc aactcaactt tcataattta aaaatctcca ttaaagttat aactagtaat   4580

tatgttttca acacttttg gttttttca ttttgttttg ctctgaccga ttcctttata   4640

tttgctcccc tatttttggc tttaatttct aattgcaaag atgtttacat caaagcttct   4700

tcacagaatt taagcaagaa atattttaat atagtgaaat ggccactact ttaagtatac   4760

aatctttaaa ataagaaagg gaggctaata tttttcatgc tatcaaatta tcttcaccct   4820

catcctttac atttttcaac atttttttt ctccataaat gacactactt gataggccgt   4880

tggttgtctg aagagtagaa gggaaactaa gagacagttc tctgtggttc aggaaaacta   4940

ctgatacttt caggggtggc ccaatgaggg aatccattga actggaagaa acacactgga   5000

ttgggtatgt ctacctggca gatactcaga aatgtagttt gcacttaagc tgtaatttta   5060

tttgttcttt ttctgaactc cattttggat tttgaatcaa gcaatatgga agcaaccagc   5120

aaattaacta atttaagtac atttttaaaa aaagagctaa gataaagact gtggaaatgc   5180

caaaccaagc aaattaggaa ccttgcaacg gtatccaggg actatgatga gaggccagca   5240
```

cattatcttc atatgtcacc tttgctacgc aaggaaattt gttcagttcg tatacttcgt 5300

aagaaggaat gcgagtaagg attggcttga attccatgga atttctagta tgagactatt 5360

tatatgaagt agaaggtaac tctttgcaca taaattggta taataaaaag aaaaacacaa 5420

acattcaaag cttagggata ggtccttggg tcaaaagttg taaataaatg tgaaacatct 5480

tctc 5484

<210> 16
<211> 1308
<212> PRT
<213> Homo sapiens

<400> 16

Met Lys Pro Ala Thr Gly Leu Trp Val Trp Val Ser Leu Leu Val Ala
1 5 10 15

Ala Gly Thr Val Gln Pro Ser Asp Ser Gln Ser Val Cys Ala Gly Thr
20 25 30

Glu Asn Lys Leu Ser Ser Leu Ser Asp Leu Glu Gln Gln Tyr Arg Ala
35 40 45

Leu Arg Lys Tyr Tyr Glu Asn Cys Glu Val Val Met Gly Asn Leu Glu
50 55 60

Ile Thr Ser Ile Glu His Asn Arg Asp Leu Ser Phe Leu Arg Ser Val
65 70 75 80

Arg Glu Val Thr Gly Tyr Val Leu Val Ala Leu Asn Gln Phe Arg Tyr
85 90 95

Leu Pro Leu Glu Asn Leu Arg Ile Ile Arg Gly Thr Lys Leu Tyr Glu
100 105 110

Asp Arg Tyr Ala Leu Ala Ile Phe Leu Asn Tyr Arg Lys Asp Gly Asn
115 120 125

Phe Gly Leu Gln Glu Leu Gly Leu Lys Asn Leu Thr Glu Ile Leu Asn
130 135 140

Gly Gly Val Tyr Val Asp Gln Asn Lys Phe Leu Cys Tyr Ala Asp Thr
145 150 155 160

Ile His Trp Gln Asp Ile Val Arg Asn Pro Trp Pro Ser Asn Leu Thr
165 170 175

Leu Val Ser Thr Asn Gly Ser Ser Gly Cys Gly Arg Cys His Lys Ser

180 185 190

Cys Thr Gly Arg Cys Trp Gly Pro Thr Glu Asn His Cys Gln Thr Leu
195 200 205

Thr Arg Thr Val Cys Ala Glu Gln Cys Asp Gly Arg Cys Tyr Gly Pro
210 215 220

Tyr Val Ser Asp Cys Cys His Arg Glu Cys Ala Gly Gly Cys Ser Gly
225 230 235 240

Pro Lys Asp Thr Asp Cys Phe Ala Cys Met Asn Phe Asn Asp Ser Gly
245 250 255

Ala Cys Val Thr Gln Cys Pro Gln Thr Phe Val Tyr Asn Pro Thr Thr
260 265 270

Phe Gln Leu Glu His Asn Phe Asn Ala Lys Tyr Thr Tyr Gly Ala Phe
275 280 285

Cys Val Lys Lys Cys Pro His Asn Phe Val Val Asp Ser Ser Ser Cys
290 295 300

Val Arg Ala Cys Pro Ser Ser Lys Met Glu Val Glu Glu Asn Gly Ile
305 310 315 320

Lys Met Cys Lys Pro Cys Thr Asp Ile Cys Pro Lys Ala Cys Asp Gly
325 330 335

Ile Gly Thr Gly Ser Leu Met Ser Ala Gln Thr Val Asp Ser Ser Asn
340 345 350

Ile Asp Lys Phe Ile Asn Cys Thr Lys Ile Asn Gly Asn Leu Ile Phe
355 360 365

Leu Val Thr Gly Ile His Gly Asp Pro Tyr Asn Ala Ile Glu Ala Ile
370 375 380

Asp Pro Glu Lys Leu Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly
385 390 395 400

Phe Leu Asn Ile Gln Ser Trp Pro Pro Asn Met Thr Asp Phe Ser Val
405 410 415

Phe Ser Asn Leu Val Thr Ile Gly Gly Arg Val Leu Tyr Ser Gly Leu
420 425 430

Ser Leu Leu Ile Leu Lys Gln Gln Gly Ile Thr Ser Leu Gln Phe Gln

435 440 445

Ser Leu Lys Glu Ile Ser Ala Gly Asn Ile Tyr Ile Thr Asp Asn Ser
450 455 460

Asn Leu Cys Tyr Tyr His Thr Ile Asn Trp Thr Thr Leu Phe Ser Thr
465 470 475 480

Ile Asn Gln Arg Ile Val Ile Arg Asp Asn Arg Lys Ala Glu Asn Cys
485 490 495

Thr Ala Glu Gly Met Val Cys Asn His Leu Cys Ser Ser Asp Gly Cys
500 505 510

Trp Gly Pro Gly Pro Asp Gln Cys Leu Ser Cys Arg Arg Phe Ser Arg
515 520 525

Gly Arg Ile Cys Ile Glu Ser Cys Asn Leu Tyr Asp Gly Glu Phe Arg
530 535 540

Glu Phe Glu Asn Gly Ser Ile Cys Val Glu Cys Asp Pro Gln Cys Glu
545 550 555 560

Lys Met Glu Asp Gly Leu Leu Thr Cys His Gly Pro Gly Pro Asp Asn
565 570 575

Cys Thr Lys Cys Ser His Phe Lys Asp Gly Pro Asn Cys Val Glu Lys
580 585 590

Cys Pro Asp Gly Leu Gln Gly Ala Asn Ser Phe Ile Phe Lys Tyr Ala
595 600 605

Asp Pro Asp Arg Glu Cys His Pro Cys His Pro Asn Cys Thr Gln Gly
610 615 620

Cys Asn Gly Pro Thr Ser His Asp Cys Ile Tyr Tyr Pro Trp Thr Gly
625 630 635 640

His Ser Thr Leu Pro Gln His Ala Arg Thr Pro Leu Ile Ala Ala Gly
645 650 655

Val Ile Gly Gly Leu Phe Ile Leu Val Ile Val Gly Leu Thr Phe Ala
660 665 670

Val Tyr Val Arg Arg Lys Ser Ile Lys Lys Lys Arg Ala Leu Arg Arg
675 680 685

Phe Leu Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Thr Ala

690 695 700

Pro Asn Gln Ala Gln Leu Arg Ile Leu Lys Glu Thr Glu Leu Lys Arg
705 710 715 720

Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile
725 730 735

Trp Val Pro Glu Gly Glu Thr Val Lys Ile Pro Val Ala Ile Lys Ile
740 745 750

Leu Asn Glu Thr Thr Gly Pro Lys Ala Asn Val Glu Phe Met Asp Glu
755 760 765

Ala Leu Ile Met Ala Ser Met Asp His Pro His Leu Val Arg Leu Leu
770 775 780

Gly Val Cys Leu Ser Pro Thr Ile Gln Leu Val Thr Gln Leu Met Pro
785 790 795 800

His Gly Cys Leu Leu Glu Tyr Val His Glu His Lys Asp Asn Ile Gly
805 810 815

Ser Gln Leu Leu Leu Asn Trp Cys Val Gln Ile Ala Lys Gly Met Met
820 825 830

Tyr Leu Glu Glu Arg Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn
835 840 845

Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly Leu
850 855 860

Ala Arg Leu Leu Glu Gly Asp Glu Lys Glu Tyr Asn Ala Asp Gly Gly
865 870 875 880

Lys Met Pro Ile Lys Trp Met Ala Leu Glu Cys Ile His Tyr Arg Lys
885 890 895

Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Ile Trp Glu
900 905 910

Leu Met Thr Phe Gly Gly Lys Pro Tyr Asp Gly Ile Pro Thr Arg Glu
915 920 925

Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile
930 935 940

Cys Thr Ile Asp Val Tyr Met Val Met Val Lys Cys Trp Met Ile Asp

945 950 955 960

Ala Asp Ser Arg Pro Lys Phe Lys Glu Leu Ala Ala Glu Phe Ser Arg
965 970 975

Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp Asp Arg
980 985 990

Met Lys Leu Pro Ser Pro Asn Asp Ser Lys Phe Phe Gln Asn Leu Leu
995 1000 1005

Asp Glu Glu Asp Leu Glu Asp Met Met Asp Ala Glu Glu Tyr Leu
1010 1015 1020

Val Pro Gln Ala Phe Asn Ile Pro Pro Pro Ile Tyr Thr Ser Arg
1025 1030 1035

Ala Arg Ile Asp Ser Asn Arg Ser Glu Ile Gly His Ser Pro Pro
1040 1045 1050

Pro Ala Tyr Thr Pro Met Ser Gly Asn Gln Phe Val Tyr Arg Asp
1055 1060 1065

Gly Gly Phe Ala Ala Glu Gln Gly Val Ser Val Pro Tyr Arg Ala
1070 1075 1080

Pro Thr Ser Thr Ile Pro Glu Ala Pro Val Ala Gln Gly Ala Thr
1085 1090 1095

Ala Glu Ile Phe Asp Asp Ser Cys Cys Asn Gly Thr Leu Arg Lys
1100 1105 1110

Pro Val Ala Pro His Val Gln Glu Asp Ser Ser Thr Gln Arg Tyr
1115 1120 1125

Ser Ala Asp Pro Thr Val Phe Ala Pro Glu Arg Ser Pro Arg Gly
1130 1135 1140

Glu Leu Asp Glu Glu Gly Tyr Met Thr Pro Met Arg Asp Lys Pro
1145 1150 1155

Lys Gln Glu Tyr Leu Asn Pro Val Glu Glu Asn Pro Phe Val Ser
1160 1165 1170

Arg Arg Lys Asn Gly Asp Leu Gln Ala Leu Asp Asn Pro Glu Tyr
1175 1180 1185

His Asn Ala Ser Asn Gly Pro Pro Lys Ala Glu Asp Glu Tyr Val

1190 1195 1200

Asn Glu Pro Leu Tyr Leu Asn Thr Phe Ala Asn Thr Leu Gly Lys
1205 1210 1215

Ala Glu Tyr Leu Lys Asn Asn Ile Leu Ser Met Pro Glu Lys Ala
1220 1225 1230

Lys Lys Ala Phe Asp Asn Pro Asp Tyr Trp Asn His Ser Leu Pro
1235 1240 1245

Pro Arg Ser Thr Leu Gln His Pro Asp Tyr Leu Gln Glu Tyr Ser
1250 1255 1260

Thr Lys Tyr Phe Tyr Lys Gln Asn Gly Arg Ile Arg Pro Ile Val
1265 1270 1275

Ala Glu Asn Pro Glu Tyr Leu Ser Glu Phe Ser Leu Lys Pro Gly
1280 1285 1290

Thr Val Leu Pro Pro Pro Pro Tyr Arg His Arg Asn Thr Val Val
1295 1300 1305

<210> 17
<211> 4989
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (46)..(4149)
<223> IGF-1R coding sequence

<400> 17
ttttttttt ttttgagaaa gggaatttca tcccaaataa aagga atg aag tct ggc 57
Met Lys Ser Gly
1

tcc gga gga ggg tcc ccg acc tcg ctg tgg ggg ctc ctg ttt ctc tcc 105
Ser Gly Gly Gly Ser Pro Thr Ser Leu Trp Gly Leu Leu Phe Leu Ser
5 10 15 20

gcc gcg ctc tcg ctc tgg ccg acg agt gga gaa atc tgc ggg cca ggc 153
Ala Ala Leu Ser Leu Trp Pro Thr Ser Gly Glu Ile Cys Gly Pro Gly
25 30 35

atc gac atc cgc aac gac tat cag cag ctg aag cgc ctg gag aac tgc 201
Ile Asp Ile Arg Asn Asp Tyr Gln Gln Leu Lys Arg Leu Glu Asn Cys
40 45 50

acg gtg atc gag ggc tac ctc cac atc ctg ctc atc tcc aag gcc gag 249
Thr Val Ile Glu Gly Tyr Leu His Ile Leu Leu Ile Ser Lys Ala Glu
55 60 65

gac tac cgc agc tac cgc ttc ccc aag ctc acg gtc att acc gag tac 297

```
Asp Tyr Arg Ser Tyr Arg Phe Pro Lys Leu Thr Val Ile Thr Glu Tyr
    70                  75                  80

ttg ctg ctg ttc cga gtg gct ggc ctc gag agc ctc gga gac ctc ttc      345
Leu Leu Leu Phe Arg Val Ala Gly Leu Glu Ser Leu Gly Asp Leu Phe
85                  90                  95                  100

ccc aac ctc acg gtc atc cgc ggc tgg aaa ctc ttc tac aac tac gcc      393
Pro Asn Leu Thr Val Ile Arg Gly Trp Lys Leu Phe Tyr Asn Tyr Ala
                105                 110                 115

ctg gtc atc ttc gag atg acc aat ctc aag gat att ggg ctt tac aac      441
Leu Val Ile Phe Glu Met Thr Asn Leu Lys Asp Ile Gly Leu Tyr Asn
            120                 125                 130

ctg agg aac att act cgg ggg gcc atc agg att gag aaa aat gct gac      489
Leu Arg Asn Ile Thr Arg Gly Ala Ile Arg Ile Glu Lys Asn Ala Asp
        135                 140                 145

ctc tgt tac ctc tcc act gtg gac tgg tcc ctg atc ctg gat gcg gtg      537
Leu Cys Tyr Leu Ser Thr Val Asp Trp Ser Leu Ile Leu Asp Ala Val
    150                 155                 160

tcc aat aac tac att gtg ggg aat aag ccc cca aag gaa tgt ggg gac      585
Ser Asn Asn Tyr Ile Val Gly Asn Lys Pro Pro Lys Glu Cys Gly Asp
165                 170                 175                 180

ctg tgt cca ggg acc atg gag gag aag ccg atg tgt gag aag acc acc      633
Leu Cys Pro Gly Thr Met Glu Glu Lys Pro Met Cys Glu Lys Thr Thr
                185                 190                 195

atc aac aat gag tac aac tac cgc tgc tgg acc aca aac cgc tgc cag      681
Ile Asn Asn Glu Tyr Asn Tyr Arg Cys Trp Thr Thr Asn Arg Cys Gln
            200                 205                 210

aaa atg tgc cca agc acg tgt ggg aag cgg gcg tgc acc gag aac aat      729
Lys Met Cys Pro Ser Thr Cys Gly Lys Arg Ala Cys Thr Glu Asn Asn
        215                 220                 225

gag tgc tgc cac ccc gag tgc ctg ggc agc tgc agc gcg cct gac aac      777
Glu Cys Cys His Pro Glu Cys Leu Gly Ser Cys Ser Ala Pro Asp Asn
    230                 235                 240

gac acg gcc tgt gta gct tgc cgc cac tac tac tat gcc ggt gtc tgt      825
Asp Thr Ala Cys Val Ala Cys Arg His Tyr Tyr Tyr Ala Gly Val Cys
245                 250                 255                 260

gtg cct gcc tgc ccg ccc aac acc tac agg ttt gag ggc tgg cgc tgt      873
Val Pro Ala Cys Pro Pro Asn Thr Tyr Arg Phe Glu Gly Trp Arg Cys
                265                 270                 275

gtg gac cgt gac ttc tgc gcc aac atc ctc agc gcc gag agc agc gac      921
Val Asp Arg Asp Phe Cys Ala Asn Ile Leu Ser Ala Glu Ser Ser Asp
            280                 285                 290

tcc gag ggg ttt gtg atc cac gac ggc gag tgc atg cag gag tgc ccc      969
Ser Glu Gly Phe Val Ile His Asp Gly Glu Cys Met Gln Glu Cys Pro
        295                 300                 305

tcg ggc ttc atc cgc aac ggc agc cag agc atg tac tgc atc cct tgt     1017
Ser Gly Phe Ile Arg Asn Gly Ser Gln Ser Met Tyr Cys Ile Pro Cys
    310                 315                 320

gaa ggt cct tgc ccg aag gtc tgt gag gaa gaa aag aaa aca aag acc     1065
```

```
Glu Gly Pro Cys Pro Lys Val Cys Glu Glu Glu Lys Lys Thr Lys Thr
325                 330                 335                 340

att gat tct gtt act tct gct cag atg ctc caa gga tgc acc atc ttc     1113
Ile Asp Ser Val Thr Ser Ala Gln Met Leu Gln Gly Cys Thr Ile Phe
                345                 350                 355

aag ggc aat ttg ctc att aac atc cga cgg ggg aat aac att gct tca     1161
Lys Gly Asn Leu Leu Ile Asn Ile Arg Arg Gly Asn Asn Ile Ala Ser
            360                 365                 370

gag ctg gag aac ttc atg ggg ctc atc gag gtg gtg acg ggc tac gtg     1209
Glu Leu Glu Asn Phe Met Gly Leu Ile Glu Val Val Thr Gly Tyr Val
        375                 380                 385

aag atc cgc cat tct cat gcc ttg gtc tcc ttg tcc ttc cta aaa aac     1257
Lys Ile Arg His Ser His Ala Leu Val Ser Leu Ser Phe Leu Lys Asn
    390                 395                 400

ctt cgc ctc atc cta gga gag gag cag cta gaa ggg aat tac tcc ttc     1305
Leu Arg Leu Ile Leu Gly Glu Glu Gln Leu Glu Gly Asn Tyr Ser Phe
405                 410                 415                 420

tac gtc ctc gac aac cag aac ttg cag caa ctg tgg gac tgg gac cac     1353
Tyr Val Leu Asp Asn Gln Asn Leu Gln Gln Leu Trp Asp Trp Asp His
                425                 430                 435

cgc aac ctg acc atc aaa gca ggg aaa atg tac ttt gct ttc aat ccc     1401
Arg Asn Leu Thr Ile Lys Ala Gly Lys Met Tyr Phe Ala Phe Asn Pro
            440                 445                 450

aaa tta tgt gtt tcc gaa att tac cgc atg gag gaa gtg acg ggg act     1449
Lys Leu Cys Val Ser Glu Ile Tyr Arg Met Glu Glu Val Thr Gly Thr
        455                 460                 465

aaa ggg cgc caa agc aaa ggg gac ata aac acc agg aac aac ggg gag     1497
Lys Gly Arg Gln Ser Lys Gly Asp Ile Asn Thr Arg Asn Asn Gly Glu
    470                 475                 480

aga gcc tcc tgt gaa agt gac gtc ctg cat ttc acc tcc acc acc acg     1545
Arg Ala Ser Cys Glu Ser Asp Val Leu His Phe Thr Ser Thr Thr Thr
485                 490                 495                 500

tcg aag aat cgc atc atc ata acc tgg cac cgg tac cgg ccc cct gac     1593
Ser Lys Asn Arg Ile Ile Ile Thr Trp His Arg Tyr Arg Pro Pro Asp
                505                 510                 515

tac agg gat ctc atc agc ttc acc gtt tac tac aag gaa gca ccc ttt     1641
Tyr Arg Asp Leu Ile Ser Phe Thr Val Tyr Tyr Lys Glu Ala Pro Phe
            520                 525                 530

aag aat gtc aca gag tat gat ggg cag gat gcc tgc ggc tcc aac agc     1689
Lys Asn Val Thr Glu Tyr Asp Gly Gln Asp Ala Cys Gly Ser Asn Ser
        535                 540                 545

tgg aac atg gtg gac gtg gac ctc ccg ccc aac aag gac gtg gag ccc     1737
Trp Asn Met Val Asp Val Asp Leu Pro Pro Asn Lys Asp Val Glu Pro
    550                 555                 560

ggc atc tta cta cat ggg ctg aag ccc tgg act cag tac gcc gtt tac     1785
Gly Ile Leu Leu His Gly Leu Lys Pro Trp Thr Gln Tyr Ala Val Tyr
565                 570                 575                 580

gtc aag gct gtg acc ctc acc atg gtg gag aac gac cat atc cgt ggg     1833
```

```
Val Lys Ala Val Thr Leu Thr Met Val Glu Asn Asp His Ile Arg Gly
                585                 590                 595

gcc aag agt gag atc ttg tac att cgc acc aat gct tca gtt cct tcc     1881
Ala Lys Ser Glu Ile Leu Tyr Ile Arg Thr Asn Ala Ser Val Pro Ser
            600                 605                 610

att ccc ttg gac gtt ctt tca gca tcg aac tcc tct tct cag tta atc     1929
Ile Pro Leu Asp Val Leu Ser Ala Ser Asn Ser Ser Ser Gln Leu Ile
        615                 620                 625

gtg aag tgg aac cct ccc tct ctg ccc aac ggc aac ctg agt tac tac     1977
Val Lys Trp Asn Pro Pro Ser Leu Pro Asn Gly Asn Leu Ser Tyr Tyr
    630                 635                 640

att gtg cgc tgg cag cgg cag cct cag gac ggc tac ctt tac cgg cac     2025
Ile Val Arg Trp Gln Arg Gln Pro Gln Asp Gly Tyr Leu Tyr Arg His
645                 650                 655                 660

aat tac tgc tcc aaa gac aaa atc ccc atc agg aag tat gcc gac ggc     2073
Asn Tyr Cys Ser Lys Asp Lys Ile Pro Ile Arg Lys Tyr Ala Asp Gly
                665                 670                 675

acc atc gac att gag gag gtc aca gag aac ccc aag act gag gtg tgt     2121
Thr Ile Asp Ile Glu Glu Val Thr Glu Asn Pro Lys Thr Glu Val Cys
            680                 685                 690

ggt ggg gag aaa ggg cct tgc tgc gcc tgc ccc aaa act gaa gcc gag     2169
Gly Gly Glu Lys Gly Pro Cys Cys Ala Cys Pro Lys Thr Glu Ala Glu
        695                 700                 705

aag cag gcc gag aag gag gag gct gaa tac cgc aaa gtc ttt gag aat     2217
Lys Gln Ala Glu Lys Glu Glu Ala Glu Tyr Arg Lys Val Phe Glu Asn
    710                 715                 720

ttc ctg cac aac tcc atc ttc gtg ccc aga cct gaa agg aag cgg aga     2265
Phe Leu His Asn Ser Ile Phe Val Pro Arg Pro Glu Arg Lys Arg Arg
725                 730                 735                 740

gat gtc atg caa gtg gcc aac acc acc atg tcc agc cga agc agg aac     2313
Asp Val Met Gln Val Ala Asn Thr Thr Met Ser Ser Arg Ser Arg Asn
                745                 750                 755

acc acg gcc gca gac acc tac aac atc acc gac ccg gaa gag ctg gag     2361
Thr Thr Ala Ala Asp Thr Tyr Asn Ile Thr Asp Pro Glu Glu Leu Glu
            760                 765                 770

aca gag tac cct ttc ttt gag agc aga gtg gat aac aag gag aga act     2409
Thr Glu Tyr Pro Phe Phe Glu Ser Arg Val Asp Asn Lys Glu Arg Thr
        775                 780                 785

gtc att tct aac ctt cgg cct ttc aca ttg tac cgc atc gat atc cac     2457
Val Ile Ser Asn Leu Arg Pro Phe Thr Leu Tyr Arg Ile Asp Ile His
    790                 795                 800

agc tgc aac cac gag gct gag aag ctg ggc tgc agc gcc tcc aac ttc     2505
Ser Cys Asn His Glu Ala Glu Lys Leu Gly Cys Ser Ala Ser Asn Phe
805                 810                 815                 820

gtc ttt gca agg act atg ccc gca gaa gga gca gat gac att cct ggg     2553
Val Phe Ala Arg Thr Met Pro Ala Glu Gly Ala Asp Asp Ile Pro Gly
                825                 830                 835

cca gtg acc tgg gag cca agg cct gaa aac tcc atc ttt tta aag tgg     2601
```

```
Pro Val Thr Trp Glu Pro Arg Pro Glu Asn Ser Ile Phe Leu Lys Trp
            840                 845                 850

ccg gaa cct gag aat ccc aat gga ttg att cta atg tat gaa ata aaa     2649
Pro Glu Pro Glu Asn Pro Asn Gly Leu Ile Leu Met Tyr Glu Ile Lys
        855                 860                 865

tac gga tca caa gtt gag gat cag cga gaa tgt gtg tcc aga cag gaa     2697
Tyr Gly Ser Gln Val Glu Asp Gln Arg Glu Cys Val Ser Arg Gln Glu
    870                 875                 880

tac agg aag tat gga ggg gcc aag cta aac cgg cta aac ccg ggg aac     2745
Tyr Arg Lys Tyr Gly Gly Ala Lys Leu Asn Arg Leu Asn Pro Gly Asn
885                 890                 895                 900

tac aca gcc cgg att cag gcc aca tct ctc tct ggg aat ggg tcg tgg     2793
Tyr Thr Ala Arg Ile Gln Ala Thr Ser Leu Ser Gly Asn Gly Ser Trp
                905                 910                 915

aca gat cct gtg ttc ttc tat gtc cag gcc aaa aca gga tat gaa aac     2841
Thr Asp Pro Val Phe Phe Tyr Val Gln Ala Lys Thr Gly Tyr Glu Asn
            920                 925                 930

ttc atc cat ctg atc atc gct ctg ccc gtc gct gtc ctg ttg atc gtg     2889
Phe Ile His Leu Ile Ile Ala Leu Pro Val Ala Val Leu Leu Ile Val
        935                 940                 945

gga ggg ttg gtg att atg ctg tac gtc ttc cat aga aag aga aat aac     2937
Gly Gly Leu Val Ile Met Leu Tyr Val Phe His Arg Lys Arg Asn Asn
    950                 955                 960

agc agg ctg ggg aat gga gtg ctg tat gcc tct gtg aac ccg gag tac     2985
Ser Arg Leu Gly Asn Gly Val Leu Tyr Ala Ser Val Asn Pro Glu Tyr
965                 970                 975                 980

ttc agc gct gct gat gtg tac gtt cct gat gag tgg gag gtg gct cgg     3033
Phe Ser Ala Ala Asp Val Tyr Val Pro Asp Glu Trp Glu Val Ala Arg
                985                 990                 995

gag aag atc acc  atg agc cgg gaa ctt  ggg cag ggg tcg ttt  ggg      3078
Glu Lys Ile Thr  Met Ser Arg Glu Leu  Gly Gln Gly Ser Phe  Gly
            1000                1005                1010

atg gtc tat gaa  gga gtt gcc aag ggt  gtg gtg aaa gat gaa  cct      3123
Met Val Tyr Glu  Gly Val Ala Lys Gly  Val Val Lys Asp Glu  Pro
            1015                1020                1025

gaa acc aga gtg  gcc att aaa aca gtg  aac gag gcc gca agc  atg      3168
Glu Thr Arg Val  Ala Ile Lys Thr Val  Asn Glu Ala Ala Ser  Met
            1030                1035                1040

cgt gag agg att  gag ttt ctc aac gaa  gct tct gtg atg aag  gag      3213
Arg Glu Arg Ile  Glu Phe Leu Asn Glu  Ala Ser Val Met Lys  Glu
            1045                1050                1055

ttc aat tgt cac  cat gtg gtg cga ttg  ctg ggt gtg gtg tcc  caa      3258
Phe Asn Cys His  His Val Val Arg Leu  Leu Gly Val Val Ser  Gln
            1060                1065                1070

ggc cag cca aca  ctg gtc atc atg gaa  ctg atg aca cgg ggc  gat      3303
Gly Gln Pro Thr  Leu Val Ile Met Glu  Leu Met Thr Arg Gly  Asp
            1075                1080                1085

ctc aaa agt tat  ctc cgg tct ctg agg  cca gaa atg gag aat  aat      3348
```

```
Leu Lys Ser Tyr  Leu Arg Ser Leu Arg  Pro Glu Met Glu Asn  Asn
            1090                1095                1100

cca gtc cta gca  cct cca agc ctg agc  aag atg att cag atg  gcc      3393
Pro Val Leu Ala  Pro Pro Ser Leu Ser  Lys Met Ile Gln Met  Ala
            1105                1110                1115

gga gag att gca  gac ggc atg gca tac  ctc aac gcc aat aag  ttc      3438
Gly Glu Ile Ala  Asp Gly Met Ala Tyr  Leu Asn Ala Asn Lys  Phe
            1120                1125                1130

gtc cac aga gac  ctt gct gcc cgg aat  tgc atg gta gcc gaa  gat      3483
Val His Arg Asp  Leu Ala Ala Arg Asn  Cys Met Val Ala Glu  Asp
            1135                1140                1145

ttc aca gtc aaa  atc gga gat ttt ggt  atg acg cga gat atc  tat      3528
Phe Thr Val Lys  Ile Gly Asp Phe Gly  Met Thr Arg Asp Ile  Tyr
            1150                1155                1160

gag aca gac tat  tac cgg aaa gga ggc  aaa ggg ctg ctg ccc  gtg      3573
Glu Thr Asp Tyr  Tyr Arg Lys Gly Gly  Lys Gly Leu Leu Pro  Val
            1165                1170                1175

cgc tgg atg tct  cct gag tcc ctc aag  gat gga gtc ttc acc  act      3618
Arg Trp Met Ser  Pro Glu Ser Leu Lys  Asp Gly Val Phe Thr  Thr
            1180                1185                1190

tac tcg gac gtc  tgg tcc ttc ggg gtc  gtc ctc tgg gag atc  gcc      3663
Tyr Ser Asp Val  Trp Ser Phe Gly Val  Val Leu Trp Glu Ile  Ala
            1195                1200                1205

aca ctg gcc gag  cag ccc tac cag ggc  ttg tcc aac gag caa  gtc      3708
Thr Leu Ala Glu  Gln Pro Tyr Gln Gly  Leu Ser Asn Glu Gln  Val
            1210                1215                1220

ctt cgc ttc gtc  atg gag ggc ggc ctt  ctg gac aag cca gac  aac      3753
Leu Arg Phe Val  Met Glu Gly Gly Leu  Leu Asp Lys Pro Asp  Asn
            1225                1230                1235

tgt cct gac atg  ctg ttt gaa ctg atg  cgc atg tgc tgg cag  tat      3798
Cys Pro Asp Met  Leu Phe Glu Leu Met  Arg Met Cys Trp Gln  Tyr
            1240                1245                1250

aac ccc aag atg  agg cct tcc ttc ctg  gag atc atc agc agc  atc      3843
Asn Pro Lys Met  Arg Pro Ser Phe Leu  Glu Ile Ile Ser Ser  Ile
            1255                1260                1265

aaa gag gag atg  gag cct ggc ttc cgg  gag gtc tcc ttc tac  tac      3888
Lys Glu Glu Met  Glu Pro Gly Phe Arg  Glu Val Ser Phe Tyr  Tyr
            1270                1275                1280

agc gag gag aac  aag ctg ccc gag ccg  gag gag ctg gac ctg  gag      3933
Ser Glu Glu Asn  Lys Leu Pro Glu Pro  Glu Glu Leu Asp Leu  Glu
            1285                1290                1295

cca gag aac atg  gag agc gtc ccc ctg  gac ccc tcg gcc tcc  tcg      3978
Pro Glu Asn Met  Glu Ser Val Pro Leu  Asp Pro Ser Ala Ser  Ser
            1300                1305                1310

tcc tcc ctg cca  ctg ccc gac aga cac  tca gga cac aag gcc  gag      4023
Ser Ser Leu Pro  Leu Pro Asp Arg His  Ser Gly His Lys Ala  Glu
            1315                1320                1325

aac ggc ccc ggc  cct ggg gtg ctg gtc  ctc cgc gcc agc ttc  gac      4068
```

```
Asn Gly Pro Gly  Pro Gly Val Leu Val  Leu Arg Ala Ser Phe  Asp
            1330                1335                1340

gag aga cag cct  tac gcc cac atg aac  ggg ggc cgc aag aac  gag      4113
Glu Arg Gln Pro  Tyr Ala His Met Asn  Gly Gly Arg Lys Asn  Glu
            1345                1350                1355

cgg gcc ttg ccg  ctg ccc cag tct tcg  acc tgc tga tccttggatc        4159
Arg Ala Leu Pro  Leu Pro Gln Ser Ser  Thr Cys
            1360                1365

ctgaatctgt gcaaacagta acgtgtgcgc acgcgcagcg gggtgggggg ggagagagag   4219

ttttaacaat ccattcacaa gcctcctgta cctcagtgga tcttcagttc tgcccttgct   4279

gcccgcggga gacagcttct ctgcagtaaa acacatttgg gatgttcctt ttttcaatat   4339

gcaagcagct ttttattccc tgcccaaacc cttaactgac atgggccttt aagaacctta   4399

atgacaacac ttaatagcaa cagagcactt gagaaccagt ctcctcactc tgtccctgtc   4459

cttccctgtt ctccctttct ctctcctctc tgcttcataa cggaaaaata attgccacaa   4519

gtccagctgg gaagcccttt ttatcagttt gaggaagtgg ctgtccctgt ggccccatcc   4579

aaccactgta cacacccgcc tgacaccgtg ggtcattaca aaaaaacacg tggagatgga   4639

aatttttacc tttatctttc acctttctag ggacatgaaa tttacaaagg gccatcgttc   4699

atccaaggct gttaccattt taacgctgcc taattttgcc aaaatcctga actttctccc   4759

tcatcggccc ggcgctgatt cctcgtgtcc ggaggcatgg gtgagcatgg cagctggttg   4819

ctccatttga gagacacgct ggcgacacac tccgtccatc cgactgcccc tgctgtgctg   4879

ctcaaggcca caggcacaca ggtctcattg cttctgacta gattattatt tgggggaact   4939

ggacacaata ggtctttctc tcagtgaagg tggggagaag ctgaaccggc              4989
```

<210> 18
<211> 1367
<212> PRT
<213> Homo sapiens

<400> 18

```
Met Lys Ser Gly Ser Gly Gly Gly Ser Pro Thr Ser Leu Trp Gly Leu
1               5                  10                  15

Leu Phe Leu Ser Ala Ala Leu Ser Leu Trp Pro Thr Ser Gly Glu Ile
            20                  25                  30

Cys Gly Pro Gly Ile Asp Ile Arg Asn Asp Tyr Gln Gln Leu Lys Arg
        35                  40                  45

Leu Glu Asn Cys Thr Val Ile Glu Gly Tyr Leu His Ile Leu Leu Ile
    50                  55                  60

Ser Lys Ala Glu Asp Tyr Arg Ser Tyr Arg Phe Pro Lys Leu Thr Val
```

65 70 75 80

Ile Thr Glu Tyr Leu Leu Leu Phe Arg Val Ala Gly Leu Glu Ser Leu
85 90 95

Gly Asp Leu Phe Pro Asn Leu Thr Val Ile Arg Gly Trp Lys Leu Phe
100 105 110

Tyr Asn Tyr Ala Leu Val Ile Phe Glu Met Thr Asn Leu Lys Asp Ile
115 120 125

Gly Leu Tyr Asn Leu Arg Asn Ile Thr Arg Gly Ala Ile Arg Ile Glu
130 135 140

Lys Asn Ala Asp Leu Cys Tyr Leu Ser Thr Val Asp Trp Ser Leu Ile
145 150 155 160

Leu Asp Ala Val Ser Asn Asn Tyr Ile Val Gly Asn Lys Pro Pro Lys
165 170 175

Glu Cys Gly Asp Leu Cys Pro Gly Thr Met Glu Glu Lys Pro Met Cys
180 185 190

Glu Lys Thr Thr Ile Asn Asn Glu Tyr Asn Tyr Arg Cys Trp Thr Thr
195 200 205

Asn Arg Cys Gln Lys Met Cys Pro Ser Thr Cys Gly Lys Arg Ala Cys
210 215 220

Thr Glu Asn Asn Glu Cys Cys His Pro Glu Cys Leu Gly Ser Cys Ser
225 230 235 240

Ala Pro Asp Asn Asp Thr Ala Cys Val Ala Cys Arg His Tyr Tyr Tyr
245 250 255

Ala Gly Val Cys Val Pro Ala Cys Pro Pro Asn Thr Tyr Arg Phe Glu
260 265 270

Gly Trp Arg Cys Val Asp Arg Asp Phe Cys Ala Asn Ile Leu Ser Ala
275 280 285

Glu Ser Ser Asp Ser Glu Gly Phe Val Ile His Asp Gly Glu Cys Met
290 295 300

Gln Glu Cys Pro Ser Gly Phe Ile Arg Asn Gly Ser Gln Ser Met Tyr
305 310 315 320

Cys Ile Pro Cys Glu Gly Pro Cys Pro Lys Val Cys Glu Glu Glu Lys

325 330 335

Lys Thr Lys Thr Ile Asp Ser Val Thr Ser Ala Gln Met Leu Gln Gly
340 345 350

Cys Thr Ile Phe Lys Gly Asn Leu Leu Ile Asn Ile Arg Arg Gly Asn
355 360 365

Asn Ile Ala Ser Glu Leu Glu Asn Phe Met Gly Leu Ile Glu Val Val
370 375 380

Thr Gly Tyr Val Lys Ile Arg His Ser His Ala Leu Val Ser Leu Ser
385 390 395 400

Phe Leu Lys Asn Leu Arg Leu Ile Leu Gly Glu Glu Gln Leu Glu Gly
405 410 415

Asn Tyr Ser Phe Tyr Val Leu Asp Asn Gln Asn Leu Gln Gln Leu Trp
420 425 430

Asp Trp Asp His Arg Asn Leu Thr Ile Lys Ala Gly Lys Met Tyr Phe
435 440 445

Ala Phe Asn Pro Lys Leu Cys Val Ser Glu Ile Tyr Arg Met Glu Glu
450 455 460

Val Thr Gly Thr Lys Gly Arg Gln Ser Lys Gly Asp Ile Asn Thr Arg
465 470 475 480

Asn Asn Gly Glu Arg Ala Ser Cys Glu Ser Asp Val Leu His Phe Thr
485 490 495

Ser Thr Thr Thr Ser Lys Asn Arg Ile Ile Ile Thr Trp His Arg Tyr
500 505 510

Arg Pro Pro Asp Tyr Arg Asp Leu Ile Ser Phe Thr Val Tyr Tyr Lys
515 520 525

Glu Ala Pro Phe Lys Asn Val Thr Glu Tyr Asp Gly Gln Asp Ala Cys
530 535 540

Gly Ser Asn Ser Trp Asn Met Val Asp Val Asp Leu Pro Pro Asn Lys
545 550 555 560

Asp Val Glu Pro Gly Ile Leu Leu His Gly Leu Lys Pro Trp Thr Gln
565 570 575

Tyr Ala Val Tyr Val Lys Ala Val Thr Leu Thr Met Val Glu Asn Asp

580 585 590

His Ile Arg Gly Ala Lys Ser Glu Ile Leu Tyr Ile Arg Thr Asn Ala
595 600 605

Ser Val Pro Ser Ile Pro Leu Asp Val Leu Ser Ala Ser Asn Ser Ser
610 615 620

Ser Gln Leu Ile Val Lys Trp Asn Pro Pro Ser Leu Pro Asn Gly Asn
625 630 635 640

Leu Ser Tyr Tyr Ile Val Arg Trp Gln Arg Gln Pro Gln Asp Gly Tyr
645 650 655

Leu Tyr Arg His Asn Tyr Cys Ser Lys Asp Lys Ile Pro Ile Arg Lys
660 665 670

Tyr Ala Asp Gly Thr Ile Asp Ile Glu Glu Val Thr Glu Asn Pro Lys
675 680 685

Thr Glu Val Cys Gly Gly Glu Lys Gly Pro Cys Cys Ala Cys Pro Lys
690 695 700

Thr Glu Ala Glu Lys Gln Ala Glu Lys Glu Glu Ala Glu Tyr Arg Lys
705 710 715 720

Val Phe Glu Asn Phe Leu His Asn Ser Ile Phe Val Pro Arg Pro Glu
725 730 735

Arg Lys Arg Arg Asp Val Met Gln Val Ala Asn Thr Thr Met Ser Ser
740 745 750

Arg Ser Arg Asn Thr Thr Ala Ala Asp Thr Tyr Asn Ile Thr Asp Pro
755 760 765

Glu Glu Leu Glu Thr Glu Tyr Pro Phe Phe Glu Ser Arg Val Asp Asn
770 775 780

Lys Glu Arg Thr Val Ile Ser Asn Leu Arg Pro Phe Thr Leu Tyr Arg
785 790 795 800

Ile Asp Ile His Ser Cys Asn His Glu Ala Glu Lys Leu Gly Cys Ser
805 810 815

Ala Ser Asn Phe Val Phe Ala Arg Thr Met Pro Ala Glu Gly Ala Asp
820 825 830

Asp Ile Pro Gly Pro Val Thr Trp Glu Pro Arg Pro Glu Asn Ser Ile

835 840 845

Phe Leu Lys Trp Pro Glu Pro Glu Asn Pro Asn Gly Leu Ile Leu Met
850 855 860

Tyr Glu Ile Lys Tyr Gly Ser Gln Val Glu Asp Gln Arg Glu Cys Val
865 870 875 880

Ser Arg Gln Glu Tyr Arg Lys Tyr Gly Gly Ala Lys Leu Asn Arg Leu
885 890 895

Asn Pro Gly Asn Tyr Thr Ala Arg Ile Gln Ala Thr Ser Leu Ser Gly
900 905 910

Asn Gly Ser Trp Thr Asp Pro Val Phe Phe Tyr Val Gln Ala Lys Thr
915 920 925

Gly Tyr Glu Asn Phe Ile His Leu Ile Ile Ala Leu Pro Val Ala Val
930 935 940

Leu Leu Ile Val Gly Gly Leu Val Ile Met Leu Tyr Val Phe His Arg
945 950 955 960

Lys Arg Asn Asn Ser Arg Leu Gly Asn Gly Val Leu Tyr Ala Ser Val
965 970 975

Asn Pro Glu Tyr Phe Ser Ala Ala Asp Val Tyr Val Pro Asp Glu Trp
980 985 990

Glu Val Ala Arg Glu Lys Ile Thr Met Ser Arg Glu Leu Gly Gln Gly
995 1000 1005

Ser Phe Gly Met Val Tyr Glu Gly Val Ala Lys Gly Val Val Lys
1010 1015 1020

Asp Glu Pro Glu Thr Arg Val Ala Ile Lys Thr Val Asn Glu Ala
1025 1030 1035

Ala Ser Met Arg Glu Arg Ile Glu Phe Leu Asn Glu Ala Ser Val
1040 1045 1050

Met Lys Glu Phe Asn Cys His His Val Val Arg Leu Leu Gly Val
1055 1060 1065

Val Ser Gln Gly Gln Pro Thr Leu Val Ile Met Glu Leu Met Thr
1070 1075 1080

Arg Gly Asp Leu Lys Ser Tyr Leu Arg Ser Leu Arg Pro Glu Met

1085 1090 1095

Glu Asn Asn Pro Val Leu Ala Pro Pro Ser Leu Ser Lys Met Ile
1100 1105 1110

Gln Met Ala Gly Glu Ile Ala Asp Gly Met Ala Tyr Leu Asn Ala
1115 1120 1125

Asn Lys Phe Val His Arg Asp Leu Ala Ala Arg Asn Cys Met Val
1130 1135 1140

Ala Glu Asp Phe Thr Val Lys Ile Gly Asp Phe Gly Met Thr Arg
1145 1150 1155

Asp Ile Tyr Glu Thr Asp Tyr Tyr Arg Lys Gly Gly Lys Gly Leu
1160 1165 1170

Leu Pro Val Arg Trp Met Ser Pro Glu Ser Leu Lys Asp Gly Val
1175 1180 1185

Phe Thr Thr Tyr Ser Asp Val Trp Ser Phe Gly Val Val Leu Trp
1190 1195 1200

Glu Ile Ala Thr Leu Ala Glu Gln Pro Tyr Gln Gly Leu Ser Asn
1205 1210 1215

Glu Gln Val Leu Arg Phe Val Met Glu Gly Gly Leu Leu Asp Lys
1220 1225 1230

Pro Asp Asn Cys Pro Asp Met Leu Phe Glu Leu Met Arg Met Cys
1235 1240 1245

Trp Gln Tyr Asn Pro Lys Met Arg Pro Ser Phe Leu Glu Ile Ile
1250 1255 1260

Ser Ser Ile Lys Glu Glu Met Glu Pro Gly Phe Arg Glu Val Ser
1265 1270 1275

Phe Tyr Tyr Ser Glu Glu Asn Lys Leu Pro Glu Pro Glu Glu Leu
1280 1285 1290

Asp Leu Glu Pro Glu Asn Met Glu Ser Val Pro Leu Asp Pro Ser
1295 1300 1305

Ala Ser Ser Ser Ser Leu Pro Leu Pro Asp Arg His Ser Gly His
1310 1315 1320

Lys Ala Glu Asn Gly Pro Gly Pro Gly Val Leu Val Leu Arg Ala

1325 1330 1335

Ser Phe Asp Glu Arg Gln Pro Tyr Ala His Met Asn Gly Gly Arg
1340 1345 1350

Lys Asn Glu Arg Ala Leu Pro Leu Pro Gln Ser Ser Thr Cys
1355 1360 1365

<210> 19
<211> 4691
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (107)..(4255)
<223> coding sequence

<220>
<221> sig_peptide
<222> (107)..(187)
<223> signal peptide coding region

<220>
<221> misc_feature
<222> (188)..(2392)
<223> insulin receptor alpha subunit mature peptide coding region

<220>
<221> misc_feature
<222> (2393)..(4252)
<223> insulin receptor beta subunit mature peptide coding region

<400> 19
gagaaggacg cgcggccccc agcgctcttg ggggccgcct cggagcatga cccccgcggg 60

ccagcgccgc gcgcctgatc cgaggagacc ccgcgctccc gcagcc atg ggc acc 115
Met Gly Thr
1

ggg ggc cgg cgg ggg gcg gcg gcc gcg ccg ctg ctg gtg gcg gtg gcc 163
Gly Gly Arg Arg Gly Ala Ala Ala Ala Pro Leu Leu Val Ala Val Ala
5 10 15

gcg ctg cta ctg ggc gcc gcg ggc cac ctg tac ccc gga gag gtg tgt 211
Ala Leu Leu Leu Gly Ala Ala Gly His Leu Tyr Pro Gly Glu Val Cys
20 25 30 35

ccc ggc atg gat atc cgg aac aac ctc act agg ttg cat gag ctg gag 259
Pro Gly Met Asp Ile Arg Asn Asn Leu Thr Arg Leu His Glu Leu Glu
40 45 50

aat tgc tct gtc atc gaa gga cac ttg cag ata ctc ttg atg ttc aaa 307
Asn Cys Ser Val Ile Glu Gly His Leu Gln Ile Leu Leu Met Phe Lys
55 60 65

acg agg ccc gaa gat ttc cga gac ctc agt ttc ccc aaa ctc atc atg 355
Thr Arg Pro Glu Asp Phe Arg Asp Leu Ser Phe Pro Lys Leu Ile Met
70 75 80

```
atc act gat tac ttg ctg ctc ttc cgg gtc tat ggg ctc gag agc ctg      403
Ile Thr Asp Tyr Leu Leu Leu Phe Arg Val Tyr Gly Leu Glu Ser Leu
    85                  90                  95

aag gac ctg ttc ccc aac ctc acg gtc atc cgg gga tca cga ctg ttc      451
Lys Asp Leu Phe Pro Asn Leu Thr Val Ile Arg Gly Ser Arg Leu Phe
100                 105                 110                 115

ttt aac tac gcg ctg gtc atc ttc gag atg gtt cac ctc aag gaa ctc      499
Phe Asn Tyr Ala Leu Val Ile Phe Glu Met Val His Leu Lys Glu Leu
                120                 125                 130

ggc ctc tac aac ctg atg aac atc acc cgg ggt tct gtc cgc atc gag      547
Gly Leu Tyr Asn Leu Met Asn Ile Thr Arg Gly Ser Val Arg Ile Glu
            135                 140                 145

aag aac aat gag ctc tgt tac ttg gcc act atc gac tgg tcc cgt atc      595
Lys Asn Asn Glu Leu Cys Tyr Leu Ala Thr Ile Asp Trp Ser Arg Ile
        150                 155                 160

ctg gat tcc gtg gag gat aat cac atc gtg ttg aac aaa gat gac aac      643
Leu Asp Ser Val Glu Asp Asn His Ile Val Leu Asn Lys Asp Asp Asn
    165                 170                 175

gag gag tgt gga gac atc tgt ccg ggt acc gcg aag ggc aag acc aac      691
Glu Glu Cys Gly Asp Ile Cys Pro Gly Thr Ala Lys Gly Lys Thr Asn
180                 185                 190                 195

tgc ccc gcc acc gtc atc aac ggg cag ttt gtc gaa cga tgt tgg act      739
Cys Pro Ala Thr Val Ile Asn Gly Gln Phe Val Glu Arg Cys Trp Thr
                200                 205                 210

cat agt cac tgc cag aaa gtt tgc ccg acc atc tgt aag tca cac ggc      787
His Ser His Cys Gln Lys Val Cys Pro Thr Ile Cys Lys Ser His Gly
            215                 220                 225

tgc acc gcc gaa ggc ctc tgt tgc cac agc gag tgc ctg ggc aac tgt      835
Cys Thr Ala Glu Gly Leu Cys Cys His Ser Glu Cys Leu Gly Asn Cys
        230                 235                 240

tct cag ccc gac gac ccc acc aag tgc gtg gcc tgc cgc aac ttc tac      883
Ser Gln Pro Asp Asp Pro Thr Lys Cys Val Ala Cys Arg Asn Phe Tyr
    245                 250                 255

ctg gac ggc agg tgt gtg gag acc tgc ccg ccc ccg tac tac cac ttc      931
Leu Asp Gly Arg Cys Val Glu Thr Cys Pro Pro Pro Tyr Tyr His Phe
260                 265                 270                 275

cag gac tgg cgc tgt gtg aac ttc agc ttc tgc cag gac ctg cac cac      979
Gln Asp Trp Arg Cys Val Asn Phe Ser Phe Cys Gln Asp Leu His His
                280                 285                 290

aaa tgc aag aac tcg cgg agg cag ggc tgc cac caa tac gtc att cac     1027
Lys Cys Lys Asn Ser Arg Arg Gln Gly Cys His Gln Tyr Val Ile His
            295                 300                 305

aac aac aag tgc atc cct gag tgt ccc tcc ggg tac acg atg aat tcc     1075
Asn Asn Lys Cys Ile Pro Glu Cys Pro Ser Gly Tyr Thr Met Asn Ser
        310                 315                 320

agc aac ttg ctg tgc acc cca tgc ctg ggt ccc tgt ccc aag gtg tgc     1123
Ser Asn Leu Leu Cys Thr Pro Cys Leu Gly Pro Cys Pro Lys Val Cys
    325                 330                 335
```

```
cac ctc cta gaa ggc gag aag acc atc gac tcg gtg acg tct gcc cag     1171
His Leu Leu Glu Gly Glu Lys Thr Ile Asp Ser Val Thr Ser Ala Gln
340                 345                 350                 355

gag ctc cga gga tgc acc gtc atc aac ggg agt ctg atc atc aac att     1219
Glu Leu Arg Gly Cys Thr Val Ile Asn Gly Ser Leu Ile Ile Asn Ile
                360                 365                 370

cga gga ggc aac aat ctg gca gct gag cta gaa gcc aac ctc ggc ctc     1267
Arg Gly Gly Asn Asn Leu Ala Ala Glu Leu Glu Ala Asn Leu Gly Leu
            375                 380                 385

att gaa gaa att tca ggg tat cta aaa atc cgc cga tcc tac gct ctg     1315
Ile Glu Glu Ile Ser Gly Tyr Leu Lys Ile Arg Arg Ser Tyr Ala Leu
        390                 395                 400

gtg tca ctt tcc ttc ttc cgg aag tta cgt ctg att cga gga gag acc     1363
Val Ser Leu Ser Phe Phe Arg Lys Leu Arg Leu Ile Arg Gly Glu Thr
    405                 410                 415

ttg gaa att ggg aac tac tcc ttc tat gcc ttg gac aac cag aac cta     1411
Leu Glu Ile Gly Asn Tyr Ser Phe Tyr Ala Leu Asp Asn Gln Asn Leu
420                 425                 430                 435

agg cag ctc tgg gac tgg agc aaa cac aac ctc acc acc act cag ggg     1459
Arg Gln Leu Trp Asp Trp Ser Lys His Asn Leu Thr Thr Thr Gln Gly
                440                 445                 450

aaa ctc ttc ttc cac tat aac ccc aaa ctc tgc ttg tca gaa atc cac     1507
Lys Leu Phe Phe His Tyr Asn Pro Lys Leu Cys Leu Ser Glu Ile His
            455                 460                 465

aag atg gaa gaa gtt tca gga acc aag ggg cgc cag gag aga aac gac     1555
Lys Met Glu Glu Val Ser Gly Thr Lys Gly Arg Gln Glu Arg Asn Asp
        470                 475                 480

att gcc ctg aag acc aat ggg gac aag gca tcc tgt gaa aat gag tta     1603
Ile Ala Leu Lys Thr Asn Gly Asp Lys Ala Ser Cys Glu Asn Glu Leu
    485                 490                 495

ctt aaa ttt tct tac att cgg aca tct ttt gac aag atc ttg ctg aga     1651
Leu Lys Phe Ser Tyr Ile Arg Thr Ser Phe Asp Lys Ile Leu Leu Arg
500                 505                 510                 515

tgg gag ccg tac tgg ccc ccc gac ttc cga gac ctc ttg ggg ttc atg     1699
Trp Glu Pro Tyr Trp Pro Pro Asp Phe Arg Asp Leu Leu Gly Phe Met
                520                 525                 530

ctg ttc tac aaa gag gcc cct tat cag aat gtg acg gag ttc gat ggg     1747
Leu Phe Tyr Lys Glu Ala Pro Tyr Gln Asn Val Thr Glu Phe Asp Gly
            535                 540                 545

cag gat gcg tgt ggt tcc aac agt tgg acg gtg gta gac att gac cca     1795
Gln Asp Ala Cys Gly Ser Asn Ser Trp Thr Val Val Asp Ile Asp Pro
        550                 555                 560

ccc ctg agg tcc aac gac ccc aaa tca cag aac cac cca ggg tgg ctg     1843
Pro Leu Arg Ser Asn Asp Pro Lys Ser Gln Asn His Pro Gly Trp Leu
    565                 570                 575

atg cgg ggt ctc aag ccc tgg acc cag tat gcc atc ttt gtg aag acc     1891
Met Arg Gly Leu Lys Pro Trp Thr Gln Tyr Ala Ile Phe Val Lys Thr
580                 585                 590                 595
```

```
ctg gtc acc ttt tcg gat gaa cgc cgg acc tat ggg gcc aag agt gac     1939
Leu Val Thr Phe Ser Asp Glu Arg Arg Thr Tyr Gly Ala Lys Ser Asp
                600                 605                 610

atc att tat gtc cag aca gat gcc acc aac ccc tct gtg ccc ctg gat     1987
Ile Ile Tyr Val Gln Thr Asp Ala Thr Asn Pro Ser Val Pro Leu Asp
            615                 620                 625

cca atc tca gtg tct aac tca tca tcc cag att att ctg aag tgg aaa     2035
Pro Ile Ser Val Ser Asn Ser Ser Ser Gln Ile Ile Leu Lys Trp Lys
        630                 635                 640

cca ccc tcc gac ccc aat ggc aac atc acc cac tac ctg gtt ttc tgg     2083
Pro Pro Ser Asp Pro Asn Gly Asn Ile Thr His Tyr Leu Val Phe Trp
    645                 650                 655

gag agg cag gcg gaa gac agt gag ctg ttc gag ctg gat tat tgc ctc     2131
Glu Arg Gln Ala Glu Asp Ser Glu Leu Phe Glu Leu Asp Tyr Cys Leu
660                 665                 670                 675

aaa ggg ctg aag ctg ccc tcg agg acc tgg tct cca cca ttc gag tct     2179
Lys Gly Leu Lys Leu Pro Ser Arg Thr Trp Ser Pro Pro Phe Glu Ser
                680                 685                 690

gaa gat tct cag aag cac aac cag agt gag tat gag gat tcg gcc ggc     2227
Glu Asp Ser Gln Lys His Asn Gln Ser Glu Tyr Glu Asp Ser Ala Gly
            695                 700                 705

gaa tgc tgc tcc tgt cca aag aca gac tct cag atc ctg aag gag ctg     2275
Glu Cys Cys Ser Cys Pro Lys Thr Asp Ser Gln Ile Leu Lys Glu Leu
        710                 715                 720

gag gag tcc tcg ttt agg aag acg ttt gag gat tac ctg cac aac gtg     2323
Glu Glu Ser Ser Phe Arg Lys Thr Phe Glu Asp Tyr Leu His Asn Val
    725                 730                 735

gtt ttc gtc ccc aga aaa acc tct tca ggc act ggt gcc gag gac cct     2371
Val Phe Val Pro Arg Lys Thr Ser Ser Gly Thr Gly Ala Glu Asp Pro
740                 745                 750                 755

agg cca tct cgg aaa cgc agg tcc ctt ggc gat gtt ggg aat gtg acg     2419
Arg Pro Ser Arg Lys Arg Arg Ser Leu Gly Asp Val Gly Asn Val Thr
                760                 765                 770

gtg gcc gtg ccc acg gtg gca gct ttc ccc aac act tcc tcg acc agc     2467
Val Ala Val Pro Thr Val Ala Ala Phe Pro Asn Thr Ser Ser Thr Ser
            775                 780                 785

gtg ccc acg agt ccg gag gag cac agg cct ttt gag aag gtg gtg aac     2515
Val Pro Thr Ser Pro Glu Glu His Arg Pro Phe Glu Lys Val Val Asn
        790                 795                 800

aag gag tcg ctg gtc atc tcc ggc ttg cga cac ttc acg ggc tat cgc     2563
Lys Glu Ser Leu Val Ile Ser Gly Leu Arg His Phe Thr Gly Tyr Arg
    805                 810                 815

atc gag ctg cag gct tgc aac cag gac acc cct gag gaa cgg tgc agt     2611
Ile Glu Leu Gln Ala Cys Asn Gln Asp Thr Pro Glu Glu Arg Cys Ser
820                 825                 830                 835

gtg gca gcc tac gtc agt gcg agg acc atg cct gaa gcc aag gct gat     2659
Val Ala Ala Tyr Val Ser Ala Arg Thr Met Pro Glu Ala Lys Ala Asp
                840                 845                 850
```

```
gac att gtt ggc cct gtg acg cat gaa atc ttt gag aac aac gtc gtc     2707
Asp Ile Val Gly Pro Val Thr His Glu Ile Phe Glu Asn Asn Val Val
            855                 860                 865

cac ttg atg tgg cag gag ccg aag gag ccc aat ggt ctg atc gtg ctg     2755
His Leu Met Trp Gln Glu Pro Lys Glu Pro Asn Gly Leu Ile Val Leu
        870                 875                 880

tat gaa gtg agt tat cgg cga tat ggt gat gag gag ctg cat ctc tgc     2803
Tyr Glu Val Ser Tyr Arg Arg Tyr Gly Asp Glu Glu Leu His Leu Cys
    885                 890                 895

gtc tcc cgc aag cac ttc gct ctg gaa cgg ggc tgc agg ctg cgt ggg     2851
Val Ser Arg Lys His Phe Ala Leu Glu Arg Gly Cys Arg Leu Arg Gly
900                 905                 910                 915

ctg tca ccg ggg aac tac agc gtg cga atc cgg gcc acc tcc ctt gcg     2899
Leu Ser Pro Gly Asn Tyr Ser Val Arg Ile Arg Ala Thr Ser Leu Ala
                920                 925                 930

ggc aac ggc tct tgg acg gaa ccc acc tat ttc tac gtg aca gac tat     2947
Gly Asn Gly Ser Trp Thr Glu Pro Thr Tyr Phe Tyr Val Thr Asp Tyr
            935                 940                 945

tta gac gtc ccg tca aat att gca aaa att atc atc ggc ccc ctc atc     2995
Leu Asp Val Pro Ser Asn Ile Ala Lys Ile Ile Ile Gly Pro Leu Ile
        950                 955                 960

ttt gtc ttt ctc ttc agt gtt gtg att gga agt att tat cta ttc ctg     3043
Phe Val Phe Leu Phe Ser Val Val Ile Gly Ser Ile Tyr Leu Phe Leu
    965                 970                 975

aga aag agg cag cca gat ggg ccg ctg gga ccg ctt tac gct tct tca     3091
Arg Lys Arg Gln Pro Asp Gly Pro Leu Gly Pro Leu Tyr Ala Ser Ser
980                 985                 990                 995

aac cct gag tat ctc  agt gcc agt gat gtg  ttt cca tgc tct gtg       3136
Asn Pro Glu Tyr Leu  Ser Ala Ser Asp Val  Phe Pro Cys Ser Val
                1000                1005                1010

tac gtg ccg gac gag  tgg gag gtg tct cga  gag aag atc acc ctc       3181
Tyr Val Pro Asp Glu  Trp Glu Val Ser Arg  Glu Lys Ile Thr Leu
                1015                1020                1025

ctt cga gag ctg ggg  cag ggc tcc ttc ggc  atg gtg tat gag ggc       3226
Leu Arg Glu Leu Gly  Gln Gly Ser Phe Gly  Met Val Tyr Glu Gly
                1030                1035                1040

aat gcc agg gac atc  atc aag ggt gag gca  gag acc cgc gtg gcg       3271
Asn Ala Arg Asp Ile  Ile Lys Gly Glu Ala  Glu Thr Arg Val Ala
                1045                1050                1055

gtg aag acg gtc aac  gag tca gcc agt ctc  cga gag cgg att gag       3316
Val Lys Thr Val Asn  Glu Ser Ala Ser Leu  Arg Glu Arg Ile Glu
                1060                1065                1070

ttc ctc aat gag gcc  tcg gtc atg aag ggc  ttc acc tgc cat cac       3361
Phe Leu Asn Glu Ala  Ser Val Met Lys Gly  Phe Thr Cys His His
                1075                1080                1085

gtg gtg cgc ctc ctg  gga gtg gtg tcc aag  ggc cag ccc acg ctg       3406
Val Val Arg Leu Leu  Gly Val Val Ser Lys  Gly Gln Pro Thr Leu
                1090                1095                1100
```

```
gtg gtg atg gag ctg  atg gct cac gga gac  ctg aag agc tac ctc      3451
Val Val Met Glu Leu  Met Ala His Gly Asp  Leu Lys Ser Tyr Leu
                1105                1110                1115

cgt tct ctg cgg cca  gag gct gag aat aat  cct ggc cgc cct ccc      3496
Arg Ser Leu Arg Pro  Glu Ala Glu Asn Asn  Pro Gly Arg Pro Pro
                1120                1125                1130

cct acc ctt caa gag  atg att cag atg gcg  gca gag att gct gac      3541
Pro Thr Leu Gln Glu  Met Ile Gln Met Ala  Ala Glu Ile Ala Asp
                1135                1140                1145

ggg atg gcc tac ctg  aac gcc aag aag ttt  gtg cat cgg gac ctg      3586
Gly Met Ala Tyr Leu  Asn Ala Lys Lys Phe  Val His Arg Asp Leu
                1150                1155                1160

gca gcg aga aac tgc  atg gtc gcc cat gat  ttt act gtc aaa att      3631
Ala Ala Arg Asn Cys  Met Val Ala His Asp  Phe Thr Val Lys Ile
                1165                1170                1175

gga gac ttt gga atg  acc aga gac atc tat  gaa acg gat tac tac      3676
Gly Asp Phe Gly Met  Thr Arg Asp Ile Tyr  Glu Thr Asp Tyr Tyr
                1180                1185                1190

cgg aaa ggg ggc aag  ggt ctg ctc cct gta  cgg tgg atg gca ccg      3721
Arg Lys Gly Gly Lys  Gly Leu Leu Pro Val  Arg Trp Met Ala Pro
                1195                1200                1205

gag tcc ctg aag gat  ggg gtc ttc acc act  tct tct gac atg tgg      3766
Glu Ser Leu Lys Asp  Gly Val Phe Thr Thr  Ser Ser Asp Met Trp
                1210                1215                1220

tcc ttt ggc gtg gtc  ctt tgg gaa atc acc  agc ttg gca gaa cag      3811
Ser Phe Gly Val Val  Leu Trp Glu Ile Thr  Ser Leu Ala Glu Gln
                1225                1230                1235

cct tac caa ggc ctg  tct aat gaa cag gtg  ttg aaa ttt gtc atg      3856
Pro Tyr Gln Gly Leu  Ser Asn Glu Gln Val  Leu Lys Phe Val Met
                1240                1245                1250

gat gga ggg tat ctg  gat caa ccc gac aac  tgt cca gag aga gtc      3901
Asp Gly Gly Tyr Leu  Asp Gln Pro Asp Asn  Cys Pro Glu Arg Val
                1255                1260                1265

act gac ctc atg cgc  atg tgc tgg caa ttc  aac ccc aag atg agg      3946
Thr Asp Leu Met Arg  Met Cys Trp Gln Phe  Asn Pro Lys Met Arg
                1270                1275                1280

cca acc ttc ctg gag  att gtc aac ctg ctc  aag gac gac ctg cac      3991
Pro Thr Phe Leu Glu  Ile Val Asn Leu Leu  Lys Asp Asp Leu His
                1285                1290                1295

ccc agc ttt cca gag  gtg tcg ttc ttc cac  agc gag gag aac aag      4036
Pro Ser Phe Pro Glu  Val Ser Phe Phe His  Ser Glu Glu Asn Lys
                1300                1305                1310

gct ccc gag agt gag  gag ctg gag atg gag  ttt gag gac atg gag      4081
Ala Pro Glu Ser Glu  Glu Leu Glu Met Glu  Phe Glu Asp Met Glu
                1315                1320                1325

aat gtg ccc ctg gac  cgt tcc tcg cac tgt  cag agg gag gag gcg      4126
Asn Val Pro Leu Asp  Arg Ser Ser His Cys  Gln Arg Glu Glu Ala
                1330                1335                1340
```

```
ggg ggc cgg gat gga  ggg tcc tcg ctg ggt  ttc aag cgg agc tac       4171
Gly Gly Arg Asp Gly  Gly Ser Ser Leu Gly  Phe Lys Arg Ser Tyr
                1345                1350                1355

gag gaa cac atc cct  tac aca cac atg aac  gga ggc aag aaa aac       4216
Glu Glu His Ile Pro  Tyr Thr His Met Asn  Gly Gly Lys Lys Asn
                1360                1365                1370

ggg cgg att ctg acc  ttg cct cgg tcc aat  cct tcc taa cagtgcctac    4265
Gly Arg Ile Leu Thr  Leu Pro Arg Ser Asn  Pro Ser
                1375                1380

cgtggcgggg gcgggcaggg gttcccattt tcgctttcct ctggtttgaa agcctctgga   4325

aaactcagga ttctcacgac tctaccatgt ccagtggagt tcagagatcg ttcctataca   4385

tttctgttca tcttaaggtg gactcgtttg gttaccaatt taactagtcc tgcagaggat   4445

ttaactgtga acctggaggg caaggggttt ccacagttgc tgctcctttg gggcaacgac   4505

ggtttcaaac caggattttg tgtttttttcg ttccccccac ccgcccccag cagatggaaa   4565

gaaagcacct gtttttacaa attctttttt ttttttttt ttttttttt ttgctggtgt   4625

ctgagcttca gtataaaaga caaaacttcc tgtttgtgga acaaaatttc gaaagaaaaa   4685

accaaa                                                             4691
```

<210> 20
<211> 1382
<212> PRT
<213> Homo sapiens

<400> 20

```
Met Gly Thr Gly Gly Arg Arg Gly Ala Ala Ala Ala Pro Leu Leu Val
1               5                   10                  15

Ala Val Ala Ala Leu Leu Leu Gly Ala Ala Gly His Leu Tyr Pro Gly
            20                  25                  30

Glu Val Cys Pro Gly Met Asp Ile Arg Asn Asn Leu Thr Arg Leu His
        35                  40                  45

Glu Leu Glu Asn Cys Ser Val Ile Glu Gly His Leu Gln Ile Leu Leu
    50                  55                  60

Met Phe Lys Thr Arg Pro Glu Asp Phe Arg Asp Leu Ser Phe Pro Lys
65                  70                  75                  80

Leu Ile Met Ile Thr Asp Tyr Leu Leu Leu Phe Arg Val Tyr Gly Leu
                85                  90                  95

Glu Ser Leu Lys Asp Leu Phe Pro Asn Leu Thr Val Ile Arg Gly Ser
            100                 105                 110
```

```
Arg Leu Phe Phe Asn Tyr Ala Leu Val Ile Phe Glu Met Val His Leu
        115                 120                 125

Lys Glu Leu Gly Leu Tyr Asn Leu Met Asn Ile Thr Arg Gly Ser Val
    130                 135                 140

Arg Ile Glu Lys Asn Asn Glu Leu Cys Tyr Leu Ala Thr Ile Asp Trp
145                 150                 155                 160

Ser Arg Ile Leu Asp Ser Val Glu Asp Asn His Ile Val Leu Asn Lys
                165                 170                 175

Asp Asp Asn Glu Glu Cys Gly Asp Ile Cys Pro Gly Thr Ala Lys Gly
            180                 185                 190

Lys Thr Asn Cys Pro Ala Thr Val Ile Asn Gly Gln Phe Val Glu Arg
        195                 200                 205

Cys Trp Thr His Ser His Cys Gln Lys Val Cys Pro Thr Ile Cys Lys
    210                 215                 220

Ser His Gly Cys Thr Ala Glu Gly Leu Cys Cys His Ser Glu Cys Leu
225                 230                 235                 240

Gly Asn Cys Ser Gln Pro Asp Asp Pro Thr Lys Cys Val Ala Cys Arg
                245                 250                 255

Asn Phe Tyr Leu Asp Gly Arg Cys Val Glu Thr Cys Pro Pro Pro Tyr
            260                 265                 270

Tyr His Phe Gln Asp Trp Arg Cys Val Asn Phe Ser Phe Cys Gln Asp
        275                 280                 285

Leu His His Lys Cys Lys Asn Ser Arg Arg Gln Gly Cys His Gln Tyr
    290                 295                 300

Val Ile His Asn Asn Lys Cys Ile Pro Glu Cys Pro Ser Gly Tyr Thr
305                 310                 315                 320

Met Asn Ser Ser Asn Leu Leu Cys Thr Pro Cys Leu Gly Pro Cys Pro
                325                 330                 335

Lys Val Cys His Leu Leu Glu Gly Glu Lys Thr Ile Asp Ser Val Thr
            340                 345                 350

Ser Ala Gln Glu Leu Arg Gly Cys Thr Val Ile Asn Gly Ser Leu Ile
        355                 360                 365
```

```
Ile Asn Ile Arg Gly Gly Asn Asn Leu Ala Ala Glu Leu Glu Ala Asn
    370                 375                 380

Leu Gly Leu Ile Glu Glu Ile Ser Gly Tyr Leu Lys Ile Arg Arg Ser
385                 390                 395                 400

Tyr Ala Leu Val Ser Leu Ser Phe Phe Arg Lys Leu Arg Leu Ile Arg
                405                 410                 415

Gly Glu Thr Leu Glu Ile Gly Asn Tyr Ser Phe Tyr Ala Leu Asp Asn
            420                 425                 430

Gln Asn Leu Arg Gln Leu Trp Asp Trp Ser Lys His Asn Leu Thr Thr
        435                 440                 445

Thr Gln Gly Lys Leu Phe Phe His Tyr Asn Pro Lys Leu Cys Leu Ser
    450                 455                 460

Glu Ile His Lys Met Glu Glu Val Ser Gly Thr Lys Gly Arg Gln Glu
465                 470                 475                 480

Arg Asn Asp Ile Ala Leu Lys Thr Asn Gly Asp Lys Ala Ser Cys Glu
                485                 490                 495

Asn Glu Leu Leu Lys Phe Ser Tyr Ile Arg Thr Ser Phe Asp Lys Ile
            500                 505                 510

Leu Leu Arg Trp Glu Pro Tyr Trp Pro Pro Asp Phe Arg Asp Leu Leu
        515                 520                 525

Gly Phe Met Leu Phe Tyr Lys Glu Ala Pro Tyr Gln Asn Val Thr Glu
    530                 535                 540

Phe Asp Gly Gln Asp Ala Cys Gly Ser Asn Ser Trp Thr Val Val Asp
545                 550                 555                 560

Ile Asp Pro Pro Leu Arg Ser Asn Asp Pro Lys Ser Gln Asn His Pro
                565                 570                 575

Gly Trp Leu Met Arg Gly Leu Lys Pro Trp Thr Gln Tyr Ala Ile Phe
            580                 585                 590

Val Lys Thr Leu Val Thr Phe Ser Asp Glu Arg Arg Thr Tyr Gly Ala
        595                 600                 605

Lys Ser Asp Ile Ile Tyr Val Gln Thr Asp Ala Thr Asn Pro Ser Val
    610                 615                 620
```

Pro Leu Asp Pro Ile Ser Val Ser Asn Ser Ser Ser Gln Ile Ile Leu
625 630 635 640

Lys Trp Lys Pro Pro Ser Asp Pro Asn Gly Asn Ile Thr His Tyr Leu
645 650 655

Val Phe Trp Glu Arg Gln Ala Glu Asp Ser Glu Leu Phe Glu Leu Asp
660 665 670

Tyr Cys Leu Lys Gly Leu Lys Leu Pro Ser Arg Thr Trp Ser Pro Pro
675 680 685

Phe Glu Ser Glu Asp Ser Gln Lys His Asn Gln Ser Glu Tyr Glu Asp
690 695 700

Ser Ala Gly Glu Cys Cys Ser Cys Pro Lys Thr Asp Ser Gln Ile Leu
705 710 715 720

Lys Glu Leu Glu Glu Ser Ser Phe Arg Lys Thr Phe Glu Asp Tyr Leu
725 730 735

His Asn Val Val Phe Val Pro Arg Lys Thr Ser Ser Gly Thr Gly Ala
740 745 750

Glu Asp Pro Arg Pro Ser Arg Lys Arg Arg Ser Leu Gly Asp Val Gly
755 760 765

Asn Val Thr Val Ala Val Pro Thr Val Ala Ala Phe Pro Asn Thr Ser
770 775 780

Ser Thr Ser Val Pro Thr Ser Pro Glu Glu His Arg Pro Phe Glu Lys
785 790 795 800

Val Val Asn Lys Glu Ser Leu Val Ile Ser Gly Leu Arg His Phe Thr
805 810 815

Gly Tyr Arg Ile Glu Leu Gln Ala Cys Asn Gln Asp Thr Pro Glu Glu
820 825 830

Arg Cys Ser Val Ala Ala Tyr Val Ser Ala Arg Thr Met Pro Glu Ala
835 840 845

Lys Ala Asp Asp Ile Val Gly Pro Val Thr His Glu Ile Phe Glu Asn
850 855 860

Asn Val Val His Leu Met Trp Gln Glu Pro Lys Glu Pro Asn Gly Leu
865 870 875 880

Ile Val Leu Tyr Glu Val Ser Tyr Arg Arg Tyr Gly Asp Glu Glu Leu
885 890 895

His Leu Cys Val Ser Arg Lys His Phe Ala Leu Glu Arg Gly Cys Arg
900 905 910

Leu Arg Gly Leu Ser Pro Gly Asn Tyr Ser Val Arg Ile Arg Ala Thr
915 920 925

Ser Leu Ala Gly Asn Gly Ser Trp Thr Glu Pro Thr Tyr Phe Tyr Val
930 935 940

Thr Asp Tyr Leu Asp Val Pro Ser Asn Ile Ala Lys Ile Ile Ile Gly
945 950 955 960

Pro Leu Ile Phe Val Phe Leu Phe Ser Val Val Ile Gly Ser Ile Tyr
965 970 975

Leu Phe Leu Arg Lys Arg Gln Pro Asp Gly Pro Leu Gly Pro Leu Tyr
980 985 990

Ala Ser Ser Asn Pro Glu Tyr Leu Ser Ala Ser Asp Val Phe Pro Cys
995 1000 1005

Ser Val Tyr Val Pro Asp Glu Trp Glu Val Ser Arg Glu Lys Ile
1010 1015 1020

Thr Leu Leu Arg Glu Leu Gly Gln Gly Ser Phe Gly Met Val Tyr
1025 1030 1035

Glu Gly Asn Ala Arg Asp Ile Ile Lys Gly Glu Ala Glu Thr Arg
1040 1045 1050

Val Ala Val Lys Thr Val Asn Glu Ser Ala Ser Leu Arg Glu Arg
1055 1060 1065

Ile Glu Phe Leu Asn Glu Ala Ser Val Met Lys Gly Phe Thr Cys
1070 1075 1080

His His Val Val Arg Leu Leu Gly Val Val Ser Lys Gly Gln Pro
1085 1090 1095

Thr Leu Val Val Met Glu Leu Met Ala His Gly Asp Leu Lys Ser
1100 1105 1110

Tyr Leu Arg Ser Leu Arg Pro Glu Ala Glu Asn Asn Pro Gly Arg
1115 1120 1125

Pro Pro Pro Thr Leu Gln Glu Met Ile Gln Met Ala Ala Glu Ile
1130 1135 1140

Ala Asp Gly Met Ala Tyr Leu Asn Ala Lys Lys Phe Val His Arg
1145 1150 1155

Asp Leu Ala Ala Arg Asn Cys Met Val Ala His Asp Phe Thr Val
1160 1165 1170

Lys Ile Gly Asp Phe Gly Met Thr Arg Asp Ile Tyr Glu Thr Asp
1175 1180 1185

Tyr Tyr Arg Lys Gly Gly Lys Gly Leu Leu Pro Val Arg Trp Met
1190 1195 1200

Ala Pro Glu Ser Leu Lys Asp Gly Val Phe Thr Thr Ser Ser Asp
1205 1210 1215

Met Trp Ser Phe Gly Val Val Leu Trp Glu Ile Thr Ser Leu Ala
1220 1225 1230

Glu Gln Pro Tyr Gln Gly Leu Ser Asn Glu Gln Val Leu Lys Phe
1235 1240 1245

Val Met Asp Gly Gly Tyr Leu Asp Gln Pro Asp Asn Cys Pro Glu
1250 1255 1260

Arg Val Thr Asp Leu Met Arg Met Cys Trp Gln Phe Asn Pro Lys
1265 1270 1275

Met Arg Pro Thr Phe Leu Glu Ile Val Asn Leu Leu Lys Asp Asp
1280 1285 1290

Leu His Pro Ser Phe Pro Glu Val Ser Phe Phe His Ser Glu Glu
1295 1300 1305

Asn Lys Ala Pro Glu Ser Glu Glu Leu Glu Met Glu Phe Glu Asp
1310 1315 1320

Met Glu Asn Val Pro Leu Asp Arg Ser Ser His Cys Gln Arg Glu
1325 1330 1335

Glu Ala Gly Gly Arg Asp Gly Gly Ser Ser Leu Gly Phe Lys Arg
1340 1345 1350

Ser Tyr Glu Glu His Ile Pro Tyr Thr His Met Asn Gly Gly Lys
1355 1360 1365

Lys Asn Gly Arg Ile Leu Thr Leu Pro Arg Ser Asn Pro Ser
1370 1375 1380

<210> 21
<211> 79
<212> PRT
<213> Homo sapiens

<400> 21

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 22
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa reflects Thr or Ser variants

<400> 22

Gly Xaa His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 23
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa reflects Leu or Pro variants

<400> 23

Gly Thr His Ser Xaa Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 24
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa reflects Pro or Leu variants

<400> 24

Gly Thr His Ser Leu Xaa Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 25
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (16)..(16)
<223> Xaa reflects Leu or Gln variants

<400> 25

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Xaa
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 26
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> Xaa reflects Arg or Cys variants

<220>
<221> misc_feature
<222> (18)..(18)
<223> Xaa can be any naturally occurring amino acid

<400> 26

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Xaa Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 27
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (18)..(18)
<223> Xaa reflects Met or Leu variants

<220>
<221> misc_feature
<222> (21)..(21)
<223> Xaa can be any naturally occurring amino acid

<400> 27

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Xaa Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 28
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> Xaa reflects Gly, Asp, Ala or Val variants

<220>
<221> misc_feature
<222> (36)..(36)
<223> Xaa can be any naturally occurring amino acid

<400> 28

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 29
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (31)..(31)
<223> Xaa reflects Arg or Ile variants

<220>
<221> misc_feature
<222> (54)..(54)
<223> Xaa can be any naturally occurring amino acid

<400> 29

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val Gly Arg Gly Pro
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 30
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (36)..(36)
<223> Xaa reflects Leu or Ile variants

<220>
<221> misc_feature
<222> (64)..(64)
<223> Xaa can be any naturally occurring amino acid

<400> 30

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Xaa
50 55 60

Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 31
<211> 79
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (54)..(54)
<223> Xaa reflects Pro or Arg variants

<220>
<221> misc_feature
<222> (73)..(73)
<223> Xaa can be any naturally occurring amino acid

<400> 31

Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val Pro Val Pro Leu
1 5 10 15

Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser Phe Leu Arg Pro
20 25 30

Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro Leu Ala Pro Leu
35 40 45

Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val Gly Arg Gly Pro

50 55 60

Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg Tyr Glu Gly
65 70 75

<210> 32
<211> 419
<212> PRT
<213> Homo sapiens

<400> 32

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 33
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (342)..(342)
<223> Xaa reflects Thr or Ser variants

<400> 33

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys

225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Xaa His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 34
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (345)..(345)
<223> Xaa reflects Leu or Pro variants

<400> 34

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                  10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255
```

```
His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335

Pro Cys Ala Arg Gly Thr His Ser Xaa Pro Pro Arg Pro Ala Ala Val
            340                 345                 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
        355                 360                 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
    370                 375                 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385                 390                 395                 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
                405                 410                 415

Tyr Glu Gly


<210>  35
<211>  419
<212>  PRT
<213>  Homo sapiens


<220>
<221>  MISC_FEATURE
<222>  (346)..(346)
<223>  Xaa reflects Pro or Leu variants

<400>  35

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                   10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30
```

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Xaa Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 36
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (356)..(356)
<223> Xaa reflects Leu or Gln variants

<400> 36

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

```
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320
```

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Xaa Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 37
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (357)..(357)
<223> Xaa reflects Arg or Cys variants

<220>
<221> misc_feature
<222> (358)..(358)
<223> Xaa can be any naturally occurring amino acid

<400> 37

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

```
Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320
```

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Leu Arg Xaa Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 38
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (358)..(358)
<223> Xaa reflects Met or Leu variants

<220>
<221> misc_feature
<222> (361)..(361)
<223> Xaa can be any naturally occurring amino acid

<400> 38

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

```
Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
            340                 345                 350

Pro Val Pro Leu Arg Met Gln Pro Xaa Pro Ala His Pro Val Leu Ser
        355                 360                 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
    370                 375                 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385                 390                 395                 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
                405                 410                 415

Tyr Glu Gly
```

<210> 39
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (361)..(361)
<223> Xaa reflects Gly, Asp, Ala or Val variants

<220>
<221> misc_feature
<222> (376)..(376)
<223> Xaa can be any naturally occurring amino acid

<400> 39

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                   10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80
```

```
Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335
```

```
Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
            340                 345                 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
        355                 360                 365

Phe Leu Arg Pro Ser Trp Asp Xaa Val Ser Ala Phe Tyr Ser Leu Pro
    370                 375                 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385                 390                 395                 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
                405                 410                 415

Tyr Glu Gly
```

<210> 40
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (371)..(371)
<223> Xaa reflects Arg or Ile variants

<220>
<221> misc_feature
<222> (394)..(394)
<223> Xaa can be any naturally occurring amino acid

<400> 40

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                   10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95
```

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

```
Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
        355                 360                 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
    370                 375                 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Xaa Ile Ser Pro Val Ser Val
385                 390                 395                 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
                405                 410                 415

Tyr Glu Gly
```

<210> 41
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (376)..(376)
<223> Xaa reflects Leu or Ile variants

<220>
<221> misc_feature
<222> (404)..(404)
<223> Xaa can be any naturally occurring amino acid

<400> 41

```
Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1               5                   10                  15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
            20                  25                  30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
        35                  40                  45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
    50                  55                  60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65                  70                  75                  80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
                85                  90                  95
```

```
Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
            100                 105                 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
        115                 120                 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
    130                 135                 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145                 150                 155                 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
                165                 170                 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
            180                 185                 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
        195                 200                 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
    210                 215                 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225                 230                 235                 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
                245                 250                 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
            260                 265                 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
        275                 280                 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
    290                 295                 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305                 310                 315                 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
                325                 330                 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
            340                 345                 350
```

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Xaa Asp Pro Asp Ala His Val Ala Val Asp Leu Ser Arg
405 410 415

Tyr Glu Gly

<210> 42
<211> 419
<212> PRT
<213> Homo sapiens

<220>
<221> MISC_FEATURE
<222> (394)..(394)
<223> Xaa reflects Pro or Arg variants

<220>
<221> misc_feature
<222> (413)..(413)
<223> Xaa can be any naturally occurring amino acid

<400> 42

Met Glu Leu Ala Ala Leu Cys Arg Trp Gly Leu Leu Leu Ala Leu Leu
1 5 10 15

Pro Pro Gly Ala Ala Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys
20 25 30

Leu Arg Leu Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His
35 40 45

Leu Tyr Gln Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr
50 55 60

Leu Pro Thr Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val
65 70 75 80

Gln Gly Tyr Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu
85 90 95

Gln Arg Leu Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr
100 105 110

Ala Leu Ala Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro
115 120 125

Val Thr Gly Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser
130 135 140

Leu Thr Glu Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln
145 150 155 160

Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn
165 170 175

Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys
180 185 190

His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser
195 200 205

Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys
210 215 220

Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys
225 230 235 240

Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu
245 250 255

His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val
260 265 270

Thr Tyr Asn Thr Asp Thr Phe Glu Ser Met Pro Asn Pro Glu Gly Arg
275 280 285

Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu
290 295 300

Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln
305 310 315 320

Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys
325 330 335

Pro Cys Ala Arg Gly Thr His Ser Leu Pro Pro Arg Pro Ala Ala Val
340 345 350

Pro Val Pro Leu Arg Met Gln Pro Gly Pro Ala His Pro Val Leu Ser
355 360 365

Phe Leu Arg Pro Ser Trp Asp Leu Val Ser Ala Phe Tyr Ser Leu Pro
370 375 380

Leu Ala Pro Leu Ser Pro Thr Ser Val Pro Ile Ser Pro Val Ser Val
385 390 395 400

Gly Arg Gly Pro Asp Pro Asp Ala His Val Ala Val Xaa Leu Ser Arg
405 410 415

Tyr Glu Gly